BELL TELEPHONE SYSTEM

Properties of *p-n* junction transistors

by

W. Shockley
M. Sparks
G. K. Teal

R. L. Wallace, Jr.
W. J. Pietenpol
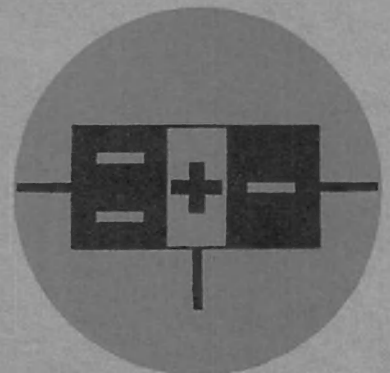
TABLE OF CONTENTS

p-n Junction Transistors

W. SHOCKLEY, M. SPARKS, AND G. K. TEAL

Bell Telephone Laboratories, Inc., Murray Hill, New Jersey

The effects of diffusion of electrons through a thin *p*-type layer of germanium have been studied in specimens consisting of two *n*-type regions with the *p*-type region interposed. It is found that potentials applied to one *n*-type region are transmitted by diffusing electrons through the *p*-type layer although the latter is grounded through an ohmic contact. When one of the *p-n* junctions is biased to saturation, power gain can be obtained through the device. Used as "*n-p-n* transistors" these units will operate on currents as low as 10 microamperes and voltages as low as 0.1 volt, have power gains of 50 db, and noise figures of about 10 db at 1000 cps. Their current-voltage characteristics are in good agreement with the diffusion theory.

I. INTRODUCTION

IN this article we shall consider the phenomena which occur when voltages are applied to a semiconductor consisting of several regions of different conductivity types. Structures of this sort, consisting in particular of two regions of one conductivity type separated by another region of the opposite type, are of great practical interest in transistor electronics. Such structures can also be used to exhibit the behavior of hole and electron diffusion in rather impressive ways. In particular, the phenomenon of "internal contact potentials" can be strikingly demonstrated with such structures.

Transistors in which the nonlinear effects originate within the germanium as a result of the relationships of *p*-type and *n*-type regions are called "*p-n* junction transistors" to distinguish them from point-contact types, in which the metal semiconductor contact often plays an essential role. There are a number of possible *p-n* junction transistor structures: The *p-n-p* transistor has been discussed previously from a theoretical viewpoint.[1,2] In this article we shall consider chiefly the *n-p-n* transistor,[3] the *n*-type phototransistor with a *p-n* hook collector and a *p-n-p-n* transistor with *p*-type emitter and *p-n* hook collector.

In the following sections we shall describe first in simple terms the basic phenomena and effects with which we are concerned. We shall next describe the actual physical structure of several of the *n-p-n* transistors and their electrical characteristics. The theoretical principles will then be put in quantitative form and the current-voltage relationships derived for certain particular models. Finally a direct comparison between theory and experiment will be presented.

II. THE *n-p-n* STRUCTURE AS A TRANSISTOR AND AS A "HOOK MULTIPLIER"

In Fig. 1 we represent an *n-p-n* structure and indicate how it may be used as a transistor. Like the type-A transistor, the current paths between emitter terminal and base and between collector terminal and base have rectifying junctions. Unlike the type-A transistor, however, the rectification arises in the interior of the germanium and not at the contacts between metal leads and the germanium; which are substantially ohmic. There are other important differences between the *n-p-n* and the type-A: In the *n-p-n* the flow of injected carriers takes place chiefly by diffusion rather than by drift in an electric field; the current multiplication at the collector, which makes possible the positive feedback instability of the type-A transistor, is lacking in the *n-p-n* transistor.

In this section we shall give a brief resumé of the theory of the operation of the transistor. Additional details will be found in the cited references. In Secs. V, *et seq.*, some aspects of the theory will be treated analytically. The analysis is simplified by the use of the following assumptions:[4]

(1) The donors and acceptors are fully ionized (this is a good assumption for germanium at room temperature).

(2) The density of minority carriers is much smaller than the density of majority carriers in each region.

(3) The net rate of recombination in any region is linear in the deviation of the minority carrier density from its thermal equilibrium value. (Assumptions (2) and (3) permit us to use linear equations in dealing with the currents arising from carrier injection.)

(4) Space charge is negligible except at the space-charge regions in the *p-n* junctions themselves.

In Fig. 1 we show the energy band diagram for the structure under consideration for zero bias and for biases applied in such a way that the unit becomes an amplifying transistor. Under the latter condition the junction J_c on the right of the figure is biased in the reverse direction. This direction is such that electrons in the *n*-type collector region have low potential energy and cannot climb the potential energy hill to the base region; similarly holes are held in the base region. Electrons in the emitter region, however, may climb the small potential hill into the base region and once in this region may diffuse so that some of them will arrive at the right-hand junction. The flow over the hill depends on the height of the hill and this height may

[1] W. Shockley, Bell System Tech. J. 28, 435–489 (1949).

[2] W. Shockley, *Electrons and Holes in Semiconductors* (D. van Nostrand Company, Inc., New York, 1950).

[3] R. L. Wallace and W. J. Pietenpol, Bell System Tech. J. July, 1951. (This article, to which we shall refer a number of times, deals with a number of practical features which we do not consider here.) It is also scheduled for the Proc. Inst. Radio Engrs. July, 1951.
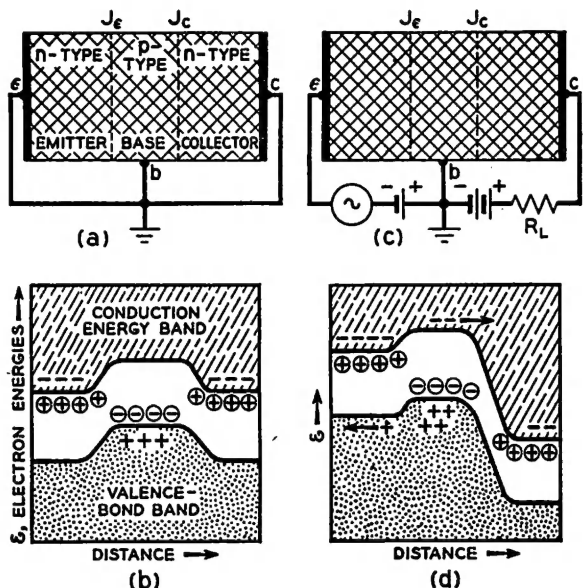
[4] These assumptions are discussed further in references 1 and 2.

FIG. 1. The *n-p-n* structure and the energy level scheme. (a) and (b) Thermal equilibrium. (c) and (d) Biased as an amplifier.

be varied by applying a variable potential to the emitter while maintaining the base at constant potential. If the base region is very thin, few of the electrons will recombine with holes in it and, as a result, efficient transmission of electron current through the layer will occur. Furthermore, if the emitter region is more highly conducting than the base region, there will be many more electrons available to climb the hill than there are holes to climb the corresponding hill in the opposite direction. As a result, most of the current across the left junction will consist of electrons. Under these conditions the behavior of this device is closely analogous to that of a vacuum tube: the emitter region corresponds to the cathode, the base to the region around the grid wires, and the collector region corresponds to the plate. In favorably designed units the controlled electron current flowing through the base region may be very much larger than the current furnished the base region for control purposes so that the transistor has a current amplification factor that is very high. It may, therefore, be operated like a grounded cathode triode with the emitter region grounded and the signal applied to the base. In Sec. IV some current-voltage characteristics for transistors are shown and for them it may be seen that the current transmission is nearly perfect.

It is interesting to note that both for the vacuum tube and for the transistor, control is accomplished by the interaction of two forms of electron flow. In the vacuum tube, metallic flow in the grid wire controls the flow of thermionic electrons in the space between grid wires. In the transistor, hole flow in the base changes the base-emitter voltage and controls electron flow through the base layer.

The structure shown in Fig. 1 with the same operating biases may be used as a collector with high multiplica-

tion in a transistor.[5] It may also be used as a phototransistor.[6] We shall consider the application as a collector in a transistor later and shall describe here how multiplication of photocurrents can occur. For this purpose there need be no electrode connected to the base layer. If light shines on the germanium near the junctions, then the hole electron pairs generated will be separated by the field of the junctions and consequently a current of holes will flow into the base layer. These holes will accumulate in the layer and will charge it positively and thus reduce its potential energy for electrons. As a result more electrons will be able to climb the hill and flow to the collector. The effect of the added holes will die away after the light is removed due to the diffusion of holes into the left region where they combine with electrons and also due to recombination with electrons which diffuse into the base layer. If the layer is very thin, however, and the density of electrons in the left region is very high, then a very large number of electrons will be able to climb over the hill for each hole that is able to enter the emitter region and recombine. In Sec. VII we shall show that the current amplification obtained in this way is proportional to the ratio of the conductivities of the two layers and is inversely proportional to the thickness of the base layer.

An interesting consequence of the diffusion of electrons through the base layer is the occurrence of "internal contact potentials."[7] In order to illustrate these we shall suppose that the base layer is grounded and that a potential is applied to the emitter. If an additional ohmic contact is also made to the base region, it will of course show ground potential. If the contact is rectifying, however, and in particular if it carries most of its current in the form of electrons, its potential will be determined by the electron density in the base layer rather than by the potential established by the grounding contact. The *n*-region on the right represents such a contact. It is found both theoretically and experimentally that if the base layer is grounded, then potentials applied to the region on the left are transmitted through the base layer, although its electrostatic potential is practically unaltered, and exhibit themselves in the region to the right, which tends to "float" (when no current is drawn from it) at a potential approximately equal to that of the left region—at least over a certain range of voltages. Theory and experiment related to this phenomenon are given in Secs. VII and VIII.

III. DESCRIPTION OF EXPERIMENTAL UNITS

The experimental units were made of a piece of single crystal germanium in which a thin *p*-type layer is

[5] W. Shockley, Phys. Rev. 78, 294 (1950).
[6] J. N. Shive who has developed the phototransistor [Phys. Rev. 76, 575 (1949)] has proposed this use of the *p-n* hook.
[7] These were first discussed in reference 1. Internal contact potentials for point contacts have been measured by G. L. Pearson and analyzed by J. Bardeen [Bell System Tech. J. 29, 469–495 (1950)].

2

interposed between two n-type sections. Units of a variety of conductivity values have been prepared. Typical values for an n-p-n structure as shown in Fig. 1 are: emitter, 100 (ohm-cm)$^{-1}$; base, 1 (ohm-cm)$^{-1}$; and collector, 0.1 (ohm-cm)$^{-1}$. Typical values of lifetime of the minority carrier in the collector section are 300–400 microseconds. The experimental evidence is that lifetimes in the other sections, in which direct measurement of lifetimes is more difficult, do not differ greatly from this. The leads to the three sections are mechanically strong and ohmic in character. The structure thus operates through conditions arising within the interior of the single crystal and not because of phenomena arising at the contacts between the leads and the germanium.

IV. PERFORMANCE CHARACTERISTICS

The performance characteristics of n-p-n transistors will be considered briefly.[8] Since these devices operate uniformly over the surfaces of p-n junctions they may be greatly altered as to the size of the active area, in contrast to point-contact devices. Thus it is possible to increase power output without corresponding increases in current density. One of the larger n-p-n transistors[9] studied as an amplifier had a junction area of 0.3 sq cm, a base layer thickness of about 0.07 cm, and delivered 2.0 watts of undistorted output in class A operation. Its frequency cutoff was about 10,000 cps, this limit being in general agreement with the effect of diffusion through the p-layer as discussed in Sec. IX.

The low power potentialities of these structures have been more thoroughly investigated. A junction area of about 0.01 sq cm and a base layer thickness of about 1.5×10^{-3} cm are typical dimensions for these units. They have operated with gains of 50 db and noise figures of about 10 db to 15 db at 1000 cycles per second. Each of these quantities is an improvement of several orders of magnitude over point-contact transistors. For low signal levels they give essentially full gain with collector voltages higher than 0.1 volt and are, therefore, exceptionally good very low power amplifiers. At some sacrifice in gain they may be operated at efficiencies of 48 to 49 percent out of a theoretical maximum of 50 percent for class A. An oscillator has been constructed by R. L. Wallace, Jr., and D. E. Thomas of this laboratory which operates on 0.6-microwatt input. The same small transistors will also operate as amplifiers with maximum output powers of several hundred milliwatts. A set of operating curves is given in Fig. 2. The frequency cutoff of these units in high gain circuits is determined chiefly by the capacitance of the collector junction and is much lower than the limit set by diffusion through the base layer. A discussion of the capaci-

[8] An extensive presentation of circuit properties is given in reference 3.

[9] The performance of this transistor was discussed at the June, 1950, Inst. Radio Engrs. Conference on Electron Devices at the University of Michigan and also at the July, 1950, Conference on Semiconductors at Reading.

tative effect will be found in the article by Wallace and Pietenpol.

V. THEORETICAL PRINCIPLES AND BOUNDARY CONDITIONS

In this and following sections we shall give analytic form to the ideas discussed in Sec. II. The principal symbols used in dealing with the theory are shown in the accompanying table of notation. In this section and the next we shall deal with the n-p-n structure in general terms and for this reason shall use subscripts "l" and "r" standing for "left" and "right" for the two n-type regions. This permits us to consider impartially cases in which either region may be biased as a collector or as an emitter.
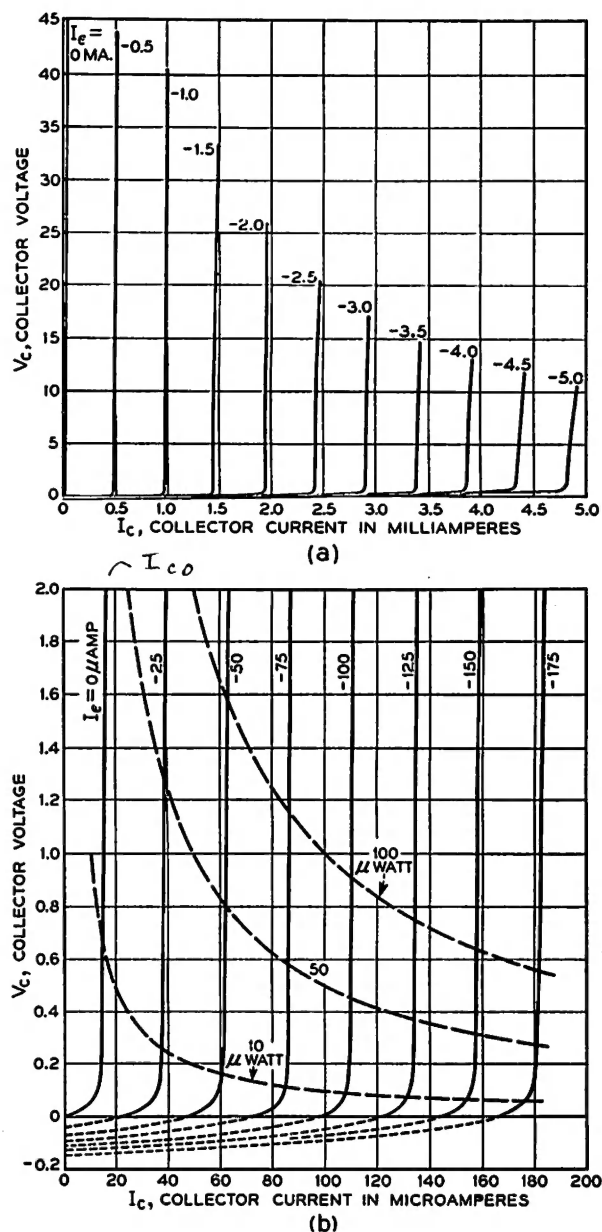

Fig. 2. Current-voltage relationships for an n-p-n transistor.

We shall first discuss the boundary conditions at the junctions when potentials are applied across them.[10] The electrostatic potential ψ in the interior of the semiconductor may have its zero chosen arbitrarily. For our purpose the zero is so chosen that $-q\psi$ is approximately the energy of an electron at a level of energy midway in the energy gap. The exact value of ψ is such that

$$n = n_i \exp q(\psi - \varphi)/kT \qquad (5.1)$$

$$p = n_i \exp q(\varphi - \psi)/kT \qquad (5.2)$$

under equilibrium conditions. Under nonequilibrium conditions, similar equations determine the "imrefs"[11] φ_n and φ_p in terms of ψ, n, and p as follows:

$$n = n_i \exp q(\psi - \varphi_n)/kT \qquad (5.3)$$

$$p = n_i \exp q(\varphi_p - \psi)/kT. \qquad (5.4)$$

NOTATION

$A =$ cross-sectional area of unit;
$b = \mu_n/\mu_p = 2.1$;
$B_l = (kT/q)[1 - \exp(-q\varphi_l/kT)]$;
$B_r = (kT/q)[1 - \exp(-q\varphi_r/kT)]$;
$e =$ base of naperian logarithms;
$D_n, D_p =$ diffusion constants for electrons and holes;
G_{ll}, G_{lr}, etc. $=$ conductances at zero voltage;
$n =$ density of electrons;
$n_b = n$ in the base layer;
$n_i = n$ in an intrinsic specimen;
$n_1 =$ deviation of n from its thermal equilibrium value;
$N_d, N_a =$ density of donors, acceptors;
$p =$ density of holes;
$q =$ charge of a hole $= -$charge of an electron;
$V, v =$ dc and ac components of voltage;
$W =$ thickness of base layer;
$\mu_n, \mu_p =$ mobilities of electrons, holes $= 3600, 1700$ cm²/volt sec;
$\sigma_i =$ intrinsic conductivity;
$\tau_{pl}, \tau_{pr}, \tau_{nb} =$ lifetimes of minority carriers;
$\varphi = $ (Fermi level)$/(-q)$;
$\varphi_n =$ imref for electrons;
$\varphi_p =$ imref for holes;
$\varphi_l, \varphi_b, \varphi_r =$ voltages of the three regions; and
$\psi =$ electrostatic potential.

The imrefs or "quasi-Fermi levels" are introduced for convenience in discussing boundary conditions at the junctions and the meaning of applied voltages. In terms of the imrefs the current densities assume a particularly simple form:

$$I_n = qD_n\nabla n + q\mu_n n E = -q\mu_n n \nabla \varphi_n \qquad (5.5)$$

$$I_p = -qD_p\nabla p + q\mu_p p E = -q\mu_p p \nabla \varphi_p. \qquad (5.6)$$

These equations show that the current densities are

---

[10] The notation used here is similar to that of references 1 and 2 and the analysis is substantially an abreviation of that of reference 1.
[11] We are indebted to the most appropriate authority for suggesting this modified name for the quasi-Fermi levels.

those corresponding to materials with the conductivity appropriate for the electron and hole densities, respectively, and to electric fields derived from the imrefs as potentials. From these relationships it is evident that a given electron current will produce much bigger changes of the imrefs when it flows in a p-type region, where the electron density is small, than it will in an n-type region. In fact the ratio of conductivity by electrons is so great between the two regions that the imref for electrons can be regarded as substantially constant in the n-type region. In accordance with the assumption that the minority carrier density is small compared to the majority density and the assumption that the space charge is negligible, it follows that the potential ψ is substantially uniform in the interior of each region also. If the contacts on regions l and r are so far from the junctions that no injected carriers reach them and are substantially ohmic, then it also follows that the imrefs for electrons in these regions are simply the voltage applied to the two contacts. In accordance with the assumption that current to the base contact is carried by holes, it also follows that the imref for holes in this region is equal to φ_b.

We shall now apply the aforementioned conclusions to the boundary condition at J_l. For simplicity we shall assume that the base is grounded so that we shall consider in general cases in which

$$\varphi_l \neq 0, \quad \varphi_b = 0, \quad \varphi_r \neq 0. \qquad (5.7)$$

By the reasoning of the preceding paragraph, the imref for electrons is continuous across the junction J_l and in fact has its largest gradient only after the interior of the base region is reached. Consequently we may take φ_n in the base region near J_l as substantially equal to φ_l. The electron density in the base region near J_l is thus given by

$$n(\text{in } b \text{ near } J_l) = n_b \exp(-q\varphi_l/kT), \qquad (5.8)$$

where n_b is the thermal equilibrium concentration of electrons at the corresponding point. (This equation follows directly from (5.3) together with the conclusion previously reached that ψ and φ_p in the base layer are unaffected by the applied potentials of (5.7).)

We shall be chiefly concerned with deviations of the densities from their equilibrium values and shall use the subscript 1 to indicate such densities. The deviation corresponding to (5.8) is

$$n_1 = n_b[\exp(-q\varphi_l/kT) - 1] \equiv -n_b q B_l/kT. \qquad (5.9)$$

In this expression we have introduced the quantity B_l defined as

$$B_l \equiv (kT/q)[1 - \exp(-q\varphi_l/kT)]. \qquad (5.10)$$

This symbol is introduced since all the currents with which we shall be concerned depend functionally upon the voltages in the form (5.10). The coefficient kT/q is introduced so that B_l has the dimensions of a voltage and for small values of φ_l, B_l is in fact approximately equal to φ_l.

4

Entirely similar reasoning leads to corresponding relationships for the hole density in the l region:

$$p(\text{in } l \text{ near } J_l) = p_l \exp(-q\varphi_l/kT) \quad (5.11)$$

$$p_1 = p_l[\exp(-q\varphi_l/kT)-1] \equiv -p_l q B_l/kT. \quad (5.12)$$

VI. THE CURRENT-VOLTAGE RELATIONSHIPS

The analysis of the last section indicates that near J_l the deviations of both hole and electron densities are proportional to B_l. These deviations lead to diffusion currents, which would vanish for the case of thermal equilibrium with $B_l=0$. As a result of the linear approximations discussed in Sec. II, these currents will be proportional to B_l. In Fig. 3 the conventions selected for signs of current are shown. We shall accordingly denote the current into region b due to hole flow across J_l as follows:

$$I_{lp} = G_{llp}B_l. \quad (6.1)$$

For the case of a uniform cross section of area A and material of uniform conductivity and uniform lifetime τ_{pl}, the value of the coefficient may be easily derived:[12]

$$G_{llp} = q\mu_p p_{nl}A/L_{pl} = \sigma_i^2 bA/(1+b)^2\sigma_l L_{pl} \quad (6.2)$$

in which the diffusion length is given by the equation,

$$L_{pl} = (D_p\tau_{pl})^{\frac{1}{2}}. \quad (6.3)$$

The last form of (6.2) expresses the conductance G_{llp} in terms of the conductivity σ_i of an intrinsic sample and the actual conductivity of the n-type region. The quantity b occurring in the equation is the ratio of mobilities:

$$b = \mu_n/\mu_p. \quad (6.4)$$

The electron current flowing across J_l can also be directly evaluated for the case of $B_r=0$. Even if no potential is applied across J_r, some of the electrons injected across J_l will arrive in the r-region. This is a consequence of the fact that the deviation n_1 is required to be zero at J_r when φ_r is zero. The electron currents across the two junctions are found to be[13]

$$I_{ln} = [(q\mu_n n_b A/L_n) \coth(W/L_n)]B_l \equiv G_{lln}B_l \quad (6.5)$$

$$I_{rn} = -[(q\mu_n n_b A/L_n) \operatorname{csch}(W/L_n)]B_l \equiv G_{rln}B_l. \quad (6.6)$$

The conductance G_{lln} may be expressed in terms of the properties of the base layer as follows:

$$G_{lln} = [\sigma_i^2 bA/(1+b^2)\sigma_b L_n] \coth(W/L_n). \quad (6.7)$$

A similar treatment for J_r leads to a corresponding set of equations. In terms of the G's and the B's the current-voltage relationship may be written as follows:

$$I_l = G_{ll}B_l + G_{lr}B_r \quad (6.8)$$

$$I_r = G_{rl}B_l + G_{rr}B_r \quad (6.9)$$

---

[12] Reference 1, Eqs. (4.20) and (4.21) or reference 2, p. 316. In a specimen of finite cross section, the recombination must be described in terms of a set of normal modes. For the cross sections of the small transistors, the lowest mode dominates and its lifetime may be used in the formulas derived for the one-dimensional case. See reference 1, Appendix V.

[13] Reference 1, Eq. (5.6), modified for electron diffusion.

$$G_{ll} = G_{lln} + G_{llp}, \quad G_{lr} = G_{lrn} \quad (6.10)$$

$$G_{rl} = G_{rln}, \quad G_{rr} = G_{rrn} + G_{rrp}. \quad (6.11)$$

For low voltages we have the approximate relationship,

$$B_l \doteq \varphi_l, \quad B_r \doteq \varphi_r \quad \text{for} \quad |\varphi_{l,r}| \ll kT/q, \quad (6.12)$$

so that the coefficients in Eqs. (6.8) and (6.9) are simply the low voltage conductance components. If the potentials consist of a bias plus a small ac component so that we may write

$$\varphi_l = V_l + v_l, \quad \varphi_r = V_r + v_r, \quad (6.13)$$

then the small signal equations become

$$i_l = g_{ll}v_l + g_{lr}v_r \quad (6.14)$$

$$i_r = g_{rl}v_l + g_{rr}v_r, \quad (6.15)$$

where the relationship between the small g's and the large G's is

$$g_{ll}/G_{ll} = g_{rl}/G_{rl} = \exp(-qV_l/kT) \quad (6.16)$$

while a similar equation applies for the other two coefficients.

From these relationships one can also derive the result that each g is proportional to the deviation of its corresponding GB term from its saturation value corresponding to $B = (kT/q)$. This may be expressed symbolically as follows:

$$g = (\text{deviation of } GB \text{ from saturation}) \cdot (q/kT). \quad (6.17)$$

For the model discussed in connection with Eqs. (6.5) and (6.6), it is evident that symmetry leads to the equation,

$$G_{lr} = G_{rl}. \quad (6.18)$$

If the conductivity varies in an unsymmetrical way, however, in the middle layer or if the lifetime is greater at one side of the layer than the other, then we cannot reach the conclusion that the two G's are equal from symmetry arguments. It can be shown, however, that it is a consequence of the linear assumptions described in Sec. II that the symmetry relationship holds no
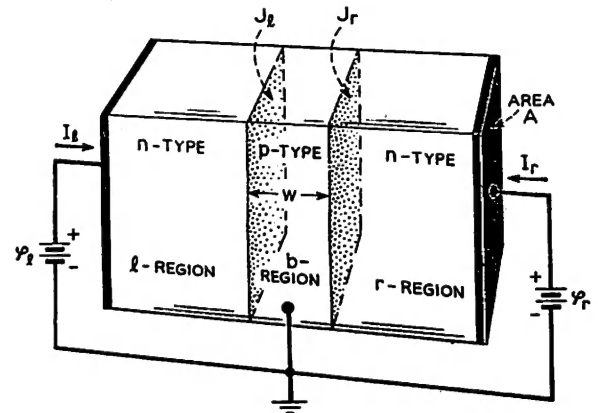

FIG. 3. Dimensions and conventions for voltage and current for an n-p-n structure.

5

matter what the geometry of the middle layer. The proof of this result is given in an appendix.[14] This leads us to the conclusion that for the idealized sort of model in which all of the currents are linear in the injected carrier density, the behavior of the transistor is described by three parameters, at least so far as low frequencies are concerned. These parameters are the four coefficients in Eqs. (6.8) and (6.9) with the relationship (6.18) between two of them.

In terms of these coefficients we may define the "alpha" of the unit using either junction as an emitter. The two α's need not be equal and the α for the left junction is given by the relationship,

$$\alpha_l \equiv -(\partial I_r/\partial I_l) \quad [\text{for} \quad \varphi_r \text{ const}] = -G_{rl}/G_{ll}$$
$$= -G_{rln}/(G_{lln}+G_{llp})$$
$$= (-G_{rln}/G_{lln})[G_{lln}/(G_{lln}+G_{llp})] = \gamma_l \beta_l. \quad (6.19)$$

The definition of γ in transistor terminology is the fraction of the current at the emitter junction produced by emitter voltage that is carried by minority carriers in the base; evidently

$$\gamma_l = G_{lln}/(G_{lln}+G_{llp}). \quad (6.20)$$

The fraction of these injected carriers that reaches the collector is defined as β:

$$\beta_l = -G_{rln}/G_{lln} = \operatorname{sech}(W/L_n). \quad (6.21)$$

The "intrinsic α" or α^* of the collector junction is defined as the ratio of change in total current per unit minority carrier current arriving at it. For a simple p-n junction collector, α^* is unity. For a "hook collector," which we treat in the next section, the arrival of injected current provokes a flow of carriers and α^* may be 100 or more.

If the middle layer is thin so that the hyberbolic functions of (6.5) and (6.6) may be approximated by their first terms, we may write

$$\beta_l \doteq 1 - \tfrac{1}{2}(W/L_n)^2 \doteq 1 \quad (6.22)$$

$$\gamma_l \doteq 1/[1 + (\sigma_b W/\sigma_l L_{pl})] \quad (6.23)$$

$$\alpha_l \doteq 1/[1 + (\sigma_b W/\sigma_l L_{pl})] \quad (6.24)$$

$$G_{lln} \doteq \sigma_1{}^2 b A/(1+b)^2 \sigma_b W. \quad (6.25)$$

Entirely similar expressions may be written for J_r simply by interchanging r and l.

[14] This symmetry result may be derived in a more general way from the reciprocity principle of electrical conduction provided no magnetic field is present. The necessary theorem is proved in section 5 of H. B. G. Casimir [Revs. Modern Phys. 17, 343 1945)]. The proof there shows that in the linear range of conductivity, G_{lr} must equal G_{rl}. Since we have shown that the currents are linear functions of the B's by an independent argument, it follows that we may take $G_{lr} = G_{rl}$ in general. The method used by Casimir is based on Onsager's principle of microscopic reversibility and has an unnecessarily abstract flavor so far as the needs of this article are concerned. The desired theorem can be proved by straightforward analytical methods as is shown in the Appendix.

VII. SPECIAL OPERATING CONDITIONS

We shall next consider the consequences of the equations derived in the previous section for several limiting cases of voltages and currents.

A. Operating as an Amplifier

In order to get into the range of linear behavior it is necessary to apply a sufficiently large reverse bias to the collector so that its current saturates. This condition corresponds to the straight parts of the characteristics shown in Fig. 2. The voltage required to get into this range must exceed about $4kT/q$. After this point is reached the emitter current and collector current may be approximated by

$$I_l = G_{ll}B_l - \alpha_l G_{ll}kT/q \quad (7.1a)$$

$$I_r = -\alpha_l G_{ll}B_l + G_{rr}kT/q$$
$$= -\alpha_l I_l + (G_{rr} + \alpha_l{}^2 G_{ll})(kT/q). \quad (7.1b)$$

Equation (7.1b) accounts for the parallel lines with currents increasing linearly in I_l shown in Fig. 2 for high collector voltages.

For applications in circuit theory, it is important to know the emitter admittance g_{ll}. To the approximation emphasized in this paper in which series ohmic resistances are neglected, this admittance is $1/r_e$ of the equivalent circuit.[3] The range of interest is usually such that $-\varphi_l > 4kT/q$ so that the application of the reasoning of (6.17) leads to

$$1/r_e = g_{ll} = \partial I_l/\partial \varphi_l = G_{ll} \exp(-q\varphi_l/kT)$$
$$\doteq I_l q/kT \doteq 40 I_l \text{mho}. \quad (7.2)$$

If the unit is operated with grounded emitter and with the collector current saturated, then the input admittance is

$$dI_b/d\varphi_b = d(-I_l - I_r)/d(-\varphi_b)$$
$$= (1 - \alpha_l)G_{ll} \exp(-q\varphi_l/kT) \quad (7.3)$$

while the transconductance is $1/(1-\alpha_l)$ larger. Thus if $\alpha_l = 0.99$, there will be a current gain of 100-fold, if the collector current is saturated. The theoretical power gain would be infinite if the collector impedance were infinite corresponding to the ideal saturation of (7.1b). Actually collector resistances of 10^7 to 10^8 ohms are observed. For currents large compared to saturation currents, (7.3) leads to an admittance of $(1-\alpha_l)I_r q/kT = 10^{-2} \times 10^{-4} \times 40 = 4 \times 10^{-5}$ mhos for $I_r = 10^{-4}$ amp, a typical value for high gain performance. The power gain will be roughly the ratio of output to input impedances times the square of the current gain and will thus be about $10^7 \times 4 \times 10^{-5} \times 10^4 \doteq 66$ db for a matched load. As discussed in Sec. IV, gains as high as 50 db have been obtained in practical circuits.

B. Hook Collector in p-n-p-n Transistor

Point contact transistors are frequently observed to have current multiplication in the sense that at a

6

fixed collector bias, changes in emitter current produce changes in collector current several fold larger. One explanation is that a "p-n hook" is formed at the collector junction so that the "intrinsic α" or α^* of the collector is high. We shall illustrate this theory for the case of an n-type transistor in which holes are injected by a p-n junction emitter and the collector consists of n and p layers.

This structure is represented in Fig. 4 and is shown for bias conditions similar to those discussed for Fig. 1 except for the added emitter region and the absence of a contact to the b layer. Regarded as a transistor, the emitter, base, and collector are ϵ', b', and c'. The operating biases put reverse voltage across J_r and forward across J_l. The base layer b will float at a potential such that the net current to it is zero:

$$I_r + I_l = G_{ll}(1-\alpha_l)B_l + G_{rr}(1-\alpha_r)B_r = 0. \quad (7.4)$$

For large reverse biases B_r will be positive and B_l negative and

$$\varphi_l - \varphi_r = V_{c'}, \quad (7.5)$$

where $V_{c'}$ is the voltage on the collector. For large $V_{c'}$, B_r will saturate and φ_l will be determined by

$$B_l = -G_{rr}(1-\alpha_r)(kT/q)/G_{ll}(1-\alpha_l) \quad (7.6)$$

so that the saturation current is

$$I_r(\text{sat.}) = (1-\alpha_r\alpha_l)G_{rr}kT/q(1-\alpha_l). \quad (7.7)$$

If holes are injected by ϵ' so that a hole current I_p^* arrives at J_r, then the condition,

$$I_r + I_l + I_p^* = 0 \quad (7.8)$$

leads to a change

$$\Delta B_l = -I_p^*/G_{ll}(1-\alpha_l) \quad (7.9)$$

and this produces an increased electron current across J_r which leads to

$$\Delta I_{nr} = G_{rl}\Delta B_l = \alpha_l I_p^*/(1-\alpha_l). \quad (7.10)$$

The "intrinsic α" or α^* of the composite structure is

$$\alpha^* = (\Delta I_{nr} + I_p^*)/I_p^* = 1/(1-\alpha_l). \quad (7.11)$$

For the thin layer approximation we see that

$$\alpha^* \doteq 1 + (\sigma_l L_{pl}/\sigma_b W). \quad (7.12)$$

Thus for thin layers and for highly conducting l regions, α^* may be made very large.

The reason for calling the structure a "p-n hook" is illustrated in Fig. 4b. The high energy potential for electrons in layer b is low potential for holes. Holes injected by ϵ' become caught in this hook and bias J_l forward so as to provoke the enhanced electron flow.

C. Phototransistor

The structure in Fig. 4 can act as a phototransistor if the hole injection by the emitter junction is simply replaced by hole electron pair generation by light. For this application, the structure has only two termi-
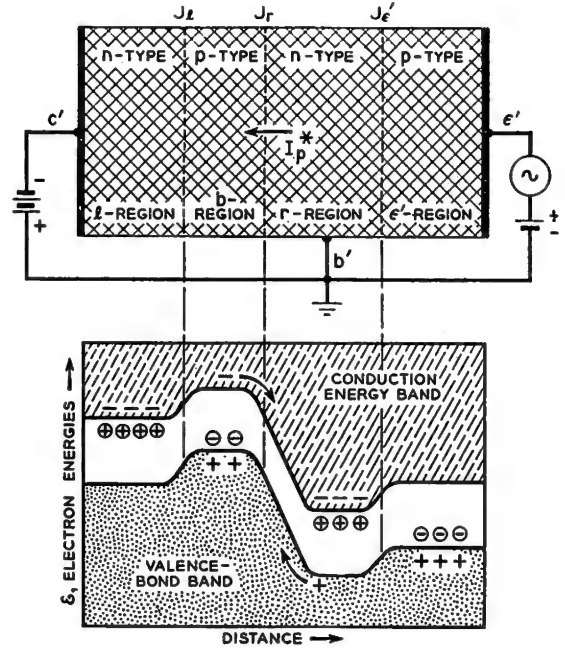

FIG. 4. A p-n-p-n transistor with a p-n hook collector.

nals, b' and c' of Fig. 4, the region ϵ' being absent. If the hole electron pairs are generated in the neighborhood of J_r, then they will be separated by the field in the junction with a current equivalent to the passage of one hole across J_r for each pair so separated. The hole current across the junction will under these conditions be multiplied just as for the case of the p-n-p-n transistor so that the apparent quantum efficiency for hole electron pairs generated at J_r is

apparent quantum efficiency

$$= \alpha^* \doteq 1 + (\sigma_l L_{pl}/\sigma_b W) \quad (7.13)$$

by the reasoning that leads to Eq. (7.12). In the following section we shall discuss some values of α^* determined by measurements of photocurrents.

D. Internal Contact Potentials

In order to discuss internal contact potentials we return to a consideration of the three region device of Fig. 3. If b is grounded and region r is allowed to seek its own potential, then potentials applied to region l will produce potentials on region r although the potential measured with an ohmic contact to b would everywhere be zero. The floating potential of region r is determined by setting the total current equal to zero so that we have

$$0 = I_r = G_{rl}B_l + G_{rr}B_r, \quad (7.14)$$

$$B_r = (-G_{rl}/G_{rr})B_l = (-G_{lr}/G_{rr})B_l = \alpha_r B_l \quad (7.15)$$

$$\exp(-q\varphi_r/kT) = 1 - \alpha_r + \alpha_r \exp(-q\varphi_l/kT). \quad (7.16)$$

Equation (7.16) expresses φ_r as a function of φ_l and α_r. It takes simple limiting form for extreme bias conditions:

7

Forward bias:

$$-q\varphi_l/kT \gg 1$$

$$\varphi_r = \varphi_l - (kT/q)\ln\alpha_r \qquad (7.17a)$$

$$\alpha_r = \exp q(\varphi_l - \varphi_r)/kT. \qquad (7.17b)$$

Zero bias:

$$|q\varphi_l/kT| \ll 1$$

$$\varphi_r = \alpha_r\varphi_l. \qquad (7.18)$$

Reverse bias:

$$q\varphi_l/kT \gg 1$$

$$\varphi_r = -(kT/q)\ln(1-\alpha_r) \qquad (7.19a)$$

$$\alpha_r = 1 - \exp(-q\varphi_r/kT). \qquad (7.19b)$$

In the following section, we shall show that these expressions are approximately satisfied.

VIII. COMPARISON WITH EXPERIMENT

In this section we shall be chiefly concerned with an analysis of data on an n-p-n transistor and with showing that it may be interpreted on the basis of the theory discussed above. The data were taken under two sets of conditions: In the first the voltages were small compared to (kT/q) and from these the G's were determined using the approximation (6.12). In the second set, a wide range of voltages were used; for all conditions, however, one of the two B's was taken as independent and the other B was either constant or else proportional to the independent B. Consequently, each current is of the form

$$I = c + mB = I_S + m[B - (kT/q)]$$
$$= I_S - I_V \exp(-qV/kT), \qquad (8.1)$$

where V is the voltage upon which B depends and

$$I_S = c + (mkT/q) \qquad (8.2)$$

is the "saturation" value of the current for large positive values of V and

$$I_V = mkT/q. \qquad (8.3)$$

Both I_S and I_V are readily calculated in terms of the G's. In the analysis of the data the measured I_S and I_V values are compared with values computed from the G's, and the dependence of I upon V is investigated.

It should be pointed out that the values of the G's are strongly dependent on temperature. Consider, for example,

$$G_{rl} = b\sigma_i^2 A/(1+b)^2\sigma_p W. \qquad (8.4)$$

TABLE 8.1. Zero bias conductances for an n-p-n transistor (conductances in micromhos at $T = 22°C$).

	Measured	Calculated
G_{11}	8.8 ± 0.5	8.8
G_{22}	33.3 ± 0.5	33.3
$G_{11}(1-\alpha_1\alpha_2)$	6.9 ± 0.5	6.9
$G_{22}(1-\alpha_1\alpha_2)$	26.5 ± 0.5	26.4
$\alpha_2 = (V_c/V_e)$ for $I_c = 0$	0.86 ± 0.02	0.89
$\alpha_1 = (V_e/V_c)$ for $I_e = 0$	0.22 ± 0.02	0.23
$G_{11} + G_{12} - 2G_{21}$	26.2 ± 0.5	26.5

In this expression[15]

$$\sigma_i^2 \propto \exp(-\epsilon_G/kT) \qquad (8.5)$$

where ϵ_G is the energy gap and

$$\sigma_p \propto T^{-\frac{3}{2}}. \qquad (8.6)$$

Consequently the value of G_{rl} at $T_0 + \Delta T = 300°K + \Delta T$ is approximately

$$G_{rl}(T_0 + \Delta T) \doteq G_{rl}(T_0)\exp[(\epsilon_G/kT_0^2) + (3/2T_0)]\Delta T$$
$$= G_{rl}(T_0)\exp(0.095 + 0.005)\Delta T \qquad (8.7)$$

so that G_{rl} increases approximately 10 percent per degree C. This increase arises chiefly from σ_i^2 and will be approximately the same for all the G's.

The fact that the G's have large temperature coefficients implies that at fixed voltages the currents will be very sensitive to temperature. This does not mean, however, that in a properly designed circuit, the behavior of the unit will be highly sensitive to temperature. The value of α_1, for example, as shown in (6.24) involves only $\sigma_b W/\sigma_e L_{pe}$ and this has only a small temperature coefficient. At a fixed emitter current, the emitter resistance is proportional to T in °K. Thus the most important quantities from a circuit point of view have small temperature coefficients.

Since in this section we are dealing with a transistor designed to have one terminal as emitter and one as collector, we shall abandon the "left" and "right" terminology and use subscript "1" for the emitter and "2" for the collector as is customary for transistors. The large signal equations are then

$$I_e = G_{11}B_1 + G_{12}B_2 \qquad (8.8)$$

$$I_c = G_{21}B_1 + G_{22}B_2 \qquad (8.9)$$

and the small signal equations are

$$i_e = g_{11}v_e + g_{12}v_c \qquad (8.10)$$

$$i_c = g_{21}v_e + g_{22}v_c \qquad (8.11)$$

where each g depends on its corresponding voltage in the form,

$$g_{ij} = G_{ij}\exp(-qV_j/kT). \qquad (8.12)$$

The low voltage conductances were measured at 2 millivolts. For this small voltage

$$B(+2\,\text{mv}) = 1.95\,\text{mv} \qquad (8.13a)$$

$$B(-2\,\text{mv}) = -2.05\,\text{mv} \qquad (8.13b)$$

so that the currents are nearly linear in this range, and the nonlinearity is practically eliminated by averaging the two polarities. With the collector grounded, I_e was measured and G_{11} computed from I_e/V_e. With the collector open circuited, I_e and V_c were measured; for this case $I_e/V_e = G_{11}(1-\alpha_1\alpha_2)$ and $V_c/V_e = -G_{21}/G_{22} = \alpha_2$. Similar data were taken with voltage applied to the

---

[15] n_i^2 varies as $T^3 \exp(-\epsilon_G/kT)$, see reference 2, page 475, and $\mu_n\mu_p$ as T^{-3}, see page 287.
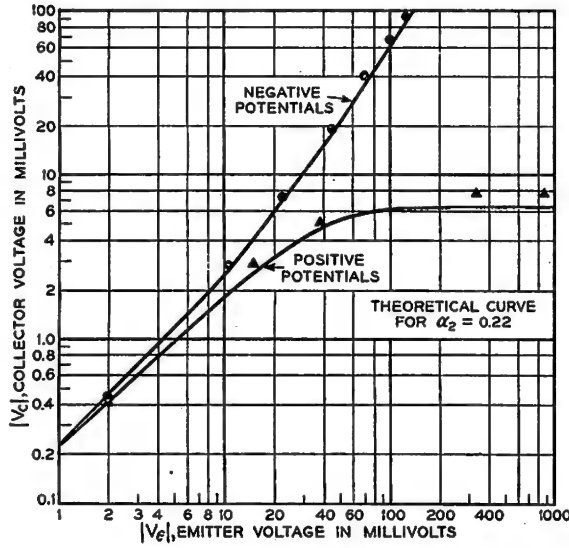
FIG. 5. Internal contact potential developed on the open circuited collector.

collector. Finally collector and emitter were connected together and the combined currents measured; this gives

$$(I_e + I_c)/V_e = G_{11} + G_{22} - G_{12} - G_{21}. \quad (8.14)$$

The values selected for the G's were: Conductances in micromhos at 22°C

$$G_{11} = 8.8, \quad G_{22} = 33.3, \quad G_{12} = G_{21} = -7.8. \quad (8.15)$$

These three values fit the seven measurements within the limits of experimental accuracy as shown in Table 8.1. The fact that the fit can be achieved with three constants is not a real test of the theory of Sec. VI, however, since any passive three terminal device in the absence of magnetic fields should satisfy the reciprocity condition and be described by three constants. What the table shows essentially is the consistency of the measurements.

Accurate values for the constants for the base layer were not available for the unit studied. However, the orders of magnitude of the G's are in reasonable agreement with values expected for a structure with constants lying in the ranges expected from the method of fabrication. We shall not attempt to obtain a perfect fit[16] but will choose as an example a structure with $A = 0.003$ cm², $W = 2.5 \times 10^{-3}$ cm, $\sigma_e = 20$, $\sigma_b = 10$, $\sigma_c = 0.1$, $\sigma_i = (1/60)$ ohm⁻¹ cm⁻¹, and lifetimes $\tau_{pc} = 40$ and $\tau_{pe} = 10$ microseconds. The resulting values for the G's based on Eqs. (6.2), (6.22), and (6.25) are

$$G_{11n} = 7.3 \text{ micromhos} \quad (8.16a)$$

$$G_{11p} = 0.5 \text{ micromhos} \quad (8.16b)$$

$$G_{22p} = 45 \text{ micromhos}. \quad (8.16c)$$

---

[16] For the simpler case of a p-n junction, the electrical properties have been predicted with an accuracy of about 20 percent from the independently measured constants describing the junction.

The value of β is greater than 0.99 for $\tau_n = 40$ microseconds and may be taken as unity so far as the G's are concerned. These values lead to

$$G_{11} = G_{11n} + G_{11p} = 7.8 \quad (8.17a)$$

$$-G_{12} = -G_{12n} = \beta_1 G_{11n} = \beta_2 G_{22n} = 7.3 \quad (8.17b)$$

$$G_{22} = G_{22n} + G_{22p} = 7.3 + 45 = 52. \quad (8.17c)$$

According to this interpretation the failure of α_1 to be unity is due chiefly to hole flow across J_e and a similar condition is true of α_2. The base current arises almost entirely from these hole flows with recombination in the base being nearly negligible.

Figures 5 and 6 show the internal contact potential effect. In Fig. 5 the potential V_e is applied and V_c measured while zero current flows to the collector terminal. The data are seen to be in general agreement with the theory and with the value of α_2, corresponding to the collector junction, obtained at low voltage for Table 8.1. Figure 6 shows similar data with the voltage applied to the collector. The data are seen to differ slightly from the theoretical curves for reverse biases. The values of α obtained by applying (7.19b) to these data are

$$\alpha_2 = 0.226 \quad \text{and} \quad \alpha_2 = 0.894, \quad (8.18)$$

which shows that the fit is very sensitive to small variations in value of α. The test of Eq. (7.17b), which applies to forward bias, cannot be carried out as satisfactorily because under these conditions the currents are relatively large and the voltage drops across the series resistances of the specimen are not negligible compared to the effects studied. The values obtained are, however, approximately the same as those of Table 8.1.

In Fig. 7, the collector is biased to saturation and I_c and I_e are plotted as functions of V_e. For this case the
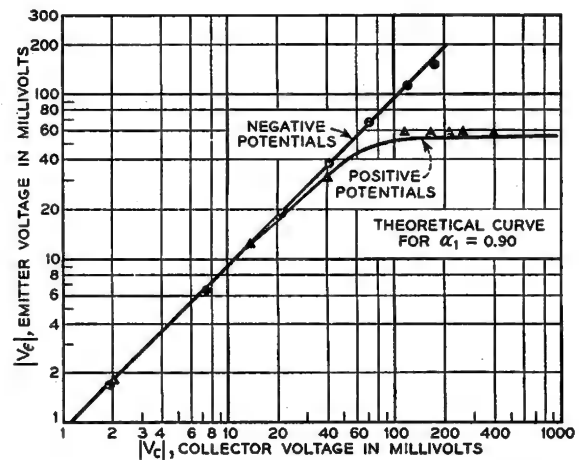

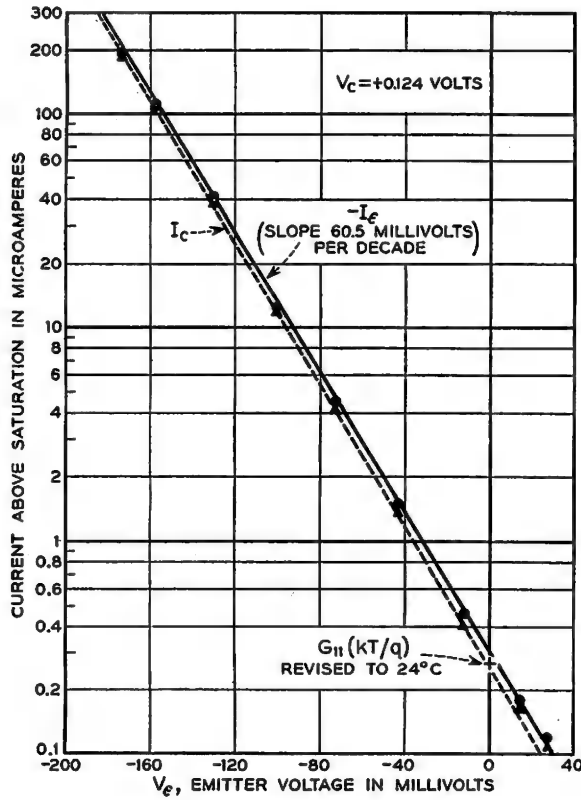FIG. 6. Internal contact potential developed on the open circuited emitter.

FIG. 7. Emitter and collector currents for the collector biased to saturation.

formulas should be

$$I_\epsilon = G_{11}B_1 + G_{12}(kT/q)$$
$$= G_{11}(1-\alpha_1)kT/q - (G_{11}kT/q)\exp(-qV_\epsilon/kT) \quad (8.19a)$$

$$I_c = G_{12}B_1 + G_{22}(kT/q)$$
$$= G_{22}(1-\alpha_2)(kT/q)$$
$$+ (\alpha_1 G_{11}kT/q)\exp(-qV_\epsilon/kT). \quad (8.19b)$$

It is seen that the lines agree well with the exponential forms and, furthermore, that the slope is in good agreement with theory which requires that for one decade of change in the current the voltage change should be

$$\Delta V = 2.30 \times kT/q = 2.30/39.4 = 59.0 \text{ mv} \quad (8.20)$$

for $T = 297°K$, the temperature at which the data were taken. The values of I_S and I_V deduced from Table 8.1 (corrected for a ΔT of 2°C) and from the data on which Fig. 7 was based are:

	Fig. 7	Table 8.1
$G_{11}(1-\alpha_1)kT/q$	0.021 μa	0.025 μa
$G_{11}(kT/q)$	0.30 μa	0.27 μa
$G_{22}(1-\alpha_2)kT/q$	0.88 μa	0.78 μa
$\alpha_1 G_{11}kT/q$	0.27 μa	0.24 μa.

The slope terms are simply the values for $V_\epsilon = 0$ in Fig. 7. The saturation values were deduced directly from the data. In the voltage range used, the satura-

tion for the collector was not perfect and the collector saturation values were corrected for a "leakage" term of about 1 megohm. (The origin of this leakage effect is not clear and it tends to saturate at higher reverse biases.)

In Fig. 8 the dependence of emitter current upon emitter voltage is again shown. The unit was at somewhat higher temperature and was measured at a higher collector voltage in one case and with the collector floating in the other. The ratio of the two terms should be

$$G_{11}/G_{11}(1-\alpha_1\alpha_2) = 1.25. \quad (8.21)$$

The observed ratio is 1.30 at $V_\epsilon = 0$ which is satisfactory agreement.

In Fig. 9 the dependence of collector current upon collector voltage is shown for two cases similar to those of Fig. 8. The ratio of the two values is again 1.30 in good agreement with the prediction. It should be noted, however, that the slope requires 74 millivolts per decade, a value appropriate to an unreasonably high temperature of 103°C. This slope is established for such low currents that it seems difficult to explain it by spurious effects of series resistances.

There is an important difference in the nature of the currents of Fig. 7, which fit the theoretical slope, and those of Fig. 9, which do not. The currents of Fig. 7 consists chiefly of electrons which diffuse through the base layer and arrive at the collector; the evidence for
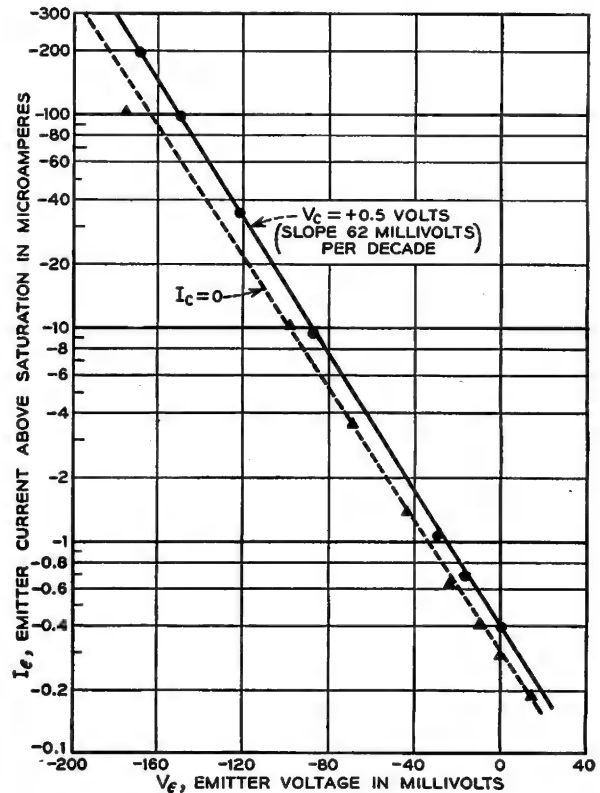

FIG. 8. Emitter current for two conditions of collectors.

10

this conclusion is the value 0.89 for α_1, which implies that only 11 percent of the current is carried by electrons recombining in the base layer or holes diffusing into and recombining in the emitter body (or the reverse processes, depending on the polarity). On the other hand, most of the collector current is probably carried by holes which recombine in the collector body. (This reasoning is in agreement with the attempt to interpret the G's in terms of the structure discussed in connection with (8.17a, b, c).) It is to be expected theoretically that if the recombination process involves trapping on recombination centers,[17] then the rate of recombination will increase less rapidly than linearly with injected carrier density because at high densities the traps tend to saturate. This view has received some experimental support from the work of F. S. Goucher and J. R. Haynes who find an apparent increase in lifetime with increasing carrier density. It may be that this mechanism accounts for failure of the currents to increase as rapidly as they should with increasing voltage in Fig. 9.

Further evidence for nonlinearity in the recombination of holes in the emitter is furnished by the dependence of α_1 upon emitter current. This can be seen in Fig. 7. As the emitter current increases, the ratio of collector to emitter current (above saturation) increases from 0.9 to about 0.95. This is interpreted as being due to the failure of hole current to increase as rapidly with voltage as does emitter current.

The tendency of α to increase with emitter current appears to be a general feature of n-p-n transistors. Wallace and Pietenpol[8] report values as high as 0.995 for α.

Phototransistors made of n-p-n structures are extremely responsive to light and exhibit apparent quantum efficiencies of at least several hundreds. In accordance with the interpretation of Sec. VII C, these efficiencies would lead to α_1 values deduced from $1-(1/\alpha^*)$ comparable to or larger than the highest observed in n-p-n transistors.

IX. SOME DESIGN CONSIDERATIONS

It has been the principal purpose of the preceding sections to examine the consequences of the diffusion theory and compare them with experiment. For this purpose the experimental conditions considered were the simplest: small currents and zero frequency. For practical applications high frequency and larger currents are also of interest. In this section we shall discuss briefly some of the factors of importance in design considerations.

Of great interest is the frequency cutoff. This may be determined by the external circuit or by one or another of several internal features of the transistor. The most fundamental of these latter is that set by the diffusion
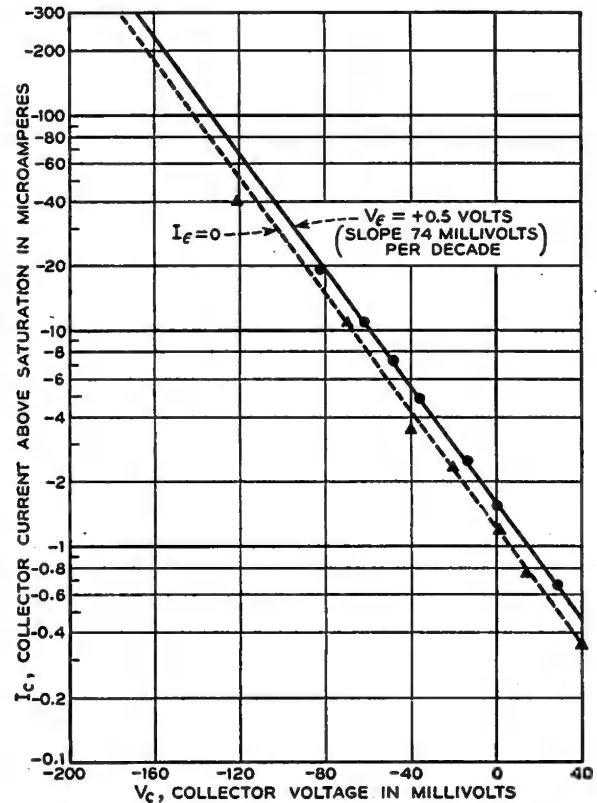

FIG. 9. Collector current for two conditions of emitter.

time through the base layer. This time is

$$\tau_D = W^2/D_n. \tag{9.1}$$

At a circular frequency ω, solutions for n_1 in the base layer are of the form,

$$\exp(i\omega t \pm x(1+i\omega\tau_n)^{\frac{1}{2}}/L_n) \tag{9.2}$$

and this leads to a β value of

$$\beta = \mathrm{sech}(1+i\omega\tau_n)^{\frac{1}{2}}W/L_n. \tag{9.3}$$

For frequencies such that $\omega\tau_n \gg 1$, this reduces to

$$B = 2/[\exp(1+i)(\omega\tau_D/2)^{\frac{1}{2}} + \exp(-1-i)(\omega\tau_D/2)^{\frac{1}{2}}]. \tag{9.4}$$

From this it is evident that for $\omega\tau_D \gg 2$ there is a phase lag of $(\omega\tau_D/2)^{\frac{1}{2}}$ radians and an equal attenuation in nepers. The power gain, which is proportional to β^2 in many cases, drops about 3 db when $\omega\tau_D = 2$ or $f = D_n/\pi W^2 \doteq 30/W^2$. For $W = 10^{-3}$ inch or 2.5×10^{-3} cm this is about 5×10^6 cps.

In addition to this fundamental limitation, there may be limitations due to capacitance and ohmic resistances. We shall illustrate this by considering the grounded emitter form of circuit, which is analogous to a grounded cathode vacuum tube circuit. For this case ac signals are applied to the base and ac voltages are developed on the collector. These voltages charge the capacitance of the base collector junction and this charging current must be furnished by hole flow in the base layer. For a layer
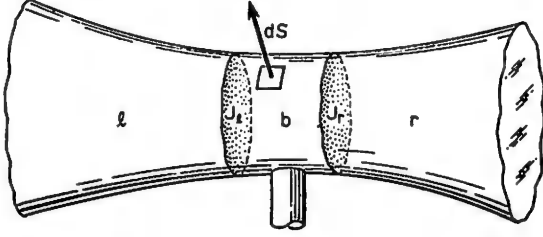
---

[17] See reference 2, page 342.

11

FIG. 10. The base region over whose surface the integration is made.

with $\sigma_b = 10$ ohm^{-1} cm^{-1} and $W = 2.5 \times 10^{-3}$ cm, the resistance from edge to opposite edge of a unit square is 40 ohms. If the capacitance is 2000 $\mu\mu f/$cm^2 and the voltage gain is 10, then the input signal will dominate the voltage on the base layer only if the width of the unit is less than

$$(10 \times \omega \times 2 \times 10^{-9} \times 40)^{-\frac{1}{2}} = (\omega 8 \times 10^{-7})^{-\frac{1}{2}}.$$

For a frequency of 5×10^6 cps, this leads to a width of 2 mm. The current through the base layer must also charge the emitter capacitance which may be much larger than the collector capacitance due to the effects of diffusion.

If the unit is too wide, the signal applied at the electrode on the base layer will be attenuated so that only a portion of the base layer will be operative in the desired way. The remainder of the base layer will be dominated by the capacitive voltages and these will induce currents which will lead to large "active" capacitances appearing between emitter and collector.

It is evident that the effects involved will lead to rather detailed calculations for any particular case but that the physical principles required to design n-p-n transistors are simply extensions of those well established for p-n junctions and n-p-n transistors for the low bias, low frequency conditions.

The low noise figures of these transistors cannot be said to be explained in the absence of an established theory of noise generation. They are, however, in rough agreement with a theory based on noise modulation of the recombination mechanism.[17] This theory predicts that each element of volume is a source of (noise current)^2 proportional to the square of the deviation of minority carrier density from its normal value. Applying this criterion to the n-p-n structure and comparing it to the type A indicates that the observed difference of 40 db or more between noise figures can be accounted for in terms of the change in current densities and geometries.

X. ACKNOWLEDGMENTS

We are indebted to a number of our colleagues for encouragement and assistance. In particular, we recognize the contribution of J. A. Morton of the transistor development group, whose encouragement resulted in the techniques that made good p-n transistors possible. We are grateful to E. Buehler and R. M. Mikulyak

APPENDIX

Proof of the Equality $G_{lr} = G_{rl}$

We shall prove this equality subject to the assumption that the electron density in the base layer is small compared to the thermal equilibrium hole density. Under these conditions, the effect of injected electrons on the potential distribution may be neglected so that the variation of the electron density from its equilibrium value may be treated by linear equations. In Fig. 10 we represent the situation considered. We shall denote by $I_{nl}(B_l, B_r)$ the current across J_l into the base carried by electrons. The symmetry relation $G_{lr} = G_{rl}$ is then established by proving that

$$I_{nl}(0, B) = I_{nr}(B, 0). \tag{A1}$$

In the base layer we shall suppose that the electrostatic potential ψ and the lifetime τ are arbitrary functions of position. The boundary condition on the external surfaces and at the metal contact will be taken as

$$I_n \cdot dS = -qn_1 s |dS|, \tag{A2}$$

where dS is the outward normal, s the surface recombination velocity and

$$n_1 = n - n_p \tag{A3}$$

is the deviation of n from the thermal equilibrium value.

We shall denote the solutions corresponding to potentials applied to the two junctions as follows:

$$B_l = B, \quad B_r = 0 \quad n_1, \quad I_n \tag{A4}$$

$$B_l = 0, \quad B_r = B \quad n_1', \quad I_n'. \tag{A5}$$

The currents in question are then

$$I_{nl}(0, B) = -\int_{J_l} I_n' \cdot dS \tag{A6}$$

$$I_{nr}(B, 0) = -\int_{J_r} I_n \cdot dS. \tag{A7}$$

The desired theorem is proved by considering the vector A:

$$A = (n_1/n_b) I_n' - (n_1'/n_b) I_n. \tag{A8}$$

Since $I_n \cdot dS$ is proportional to n_1 on the external surfaces, $A \cdot ds = 0$ on these surfaces. Hence, the integral of $A \cdot dS$ over the surface of the base region is

$$\int A \cdot dS = (qB/kT)[-I_{nl}(0, B) + I_{nr}(B, 0)] \tag{A9}$$

since on J_l and J_r we have

$$n_1/n_b = (qB/kT) \text{ on } J_l$$
$$= 0 \text{ on } J_r \tag{A10}$$

$$n_1'/n_b = 0 \text{ on } J_l$$
$$= (qB/kT) \text{ on } J_r. \tag{A11}$$

Furthermore, it can be shown that $\nabla \cdot A = 0$ in the base region and hence that (A9) is zero by Gauss's theorem so that (A1) is proved. The proof that $\nabla \cdot A = 0$ is accomplished by showing that $\nabla \cdot (n_1/n_b) I_n'$ is symmetrical in n_1 and n_1' so that the two terms in (A8) have cancelling divergences:

$$\nabla \cdot (n_1/n_b) I_n'$$
$$= (1/n_b)(\nabla n_1 - n_1 \nabla \ln n_b) \cdot I_n' + (n_1/n_b) \nabla \cdot I_n'$$
$$= (1/n_b)[(\nabla n_1 + n_1 qE/kT) \cdot (q\mu_n n_1' E + qD_n \nabla n_1')$$
$$\qquad - q(n_1 n_1'/\tau)], \tag{A12}$$

which is seen to be symmetrical.

12

Some Circuit Properties and Applications
of *n-p-n* Transistors

By R. L. WALLACE, JR. and W. J. PIETENPOL

Shockley, Sparks, and Teal have recently described the physical properties of a new kind of transistor. Preliminary studies of circuit performance show that it is a stable, high gain, quiet amplifier of considerable practical interest. This paper analyzes the performance of a few early experimental units.

INTRODUCTION

Almost two years ago, W. Shockley[1, 2] first published the theory of a transistor made from a single piece of germanium in which the conductivity type varies in such a way as to produce two rectifying junctions. Since that time, M. Sparks, G. K. Teal and others at the Bell Telephone Laboratories[3, 4] have contributed notably to the physical realization of this device.

Recently Sparks has produced a number of *n-p-n* transistors and has found their behavior to be closely in accord with Shockley's theory.[4] Preliminary circuit studies on these devices have shown that in several respects their performance is remarkable. In view of this, our transistor development group has undertaken to produce small quantities of *n-p-n* transistors in a form suitable for incorporation in working circuits.

This paper will deal principally with the circuit aspects of the *n-p-n* transistor by presenting and analyzing performance data on a small number of experimental units. For a discussion of the solid state physics of its design and operation the reader is referred to the previously mentioned works of Shockley, Sparks, and Teal.

OUTSTANDING PROPERTIES

Before getting lost in a maze of detail, it seems worthwhile to list and mention briefly the salient features of this new transistor. They are:

1. *Relatively low noise figure.* Most of the units measured so far have a noise figure between 10 and 20 db at 1000 cps.

2. *Complete freedom from short-circuit instability.* The input and output impedances are always positive whether the transistor is connected grounded-emitter, grounded-base, or grounded collector. This permits a great deal of freedom in circuit design and makes it possible, by choosing the appropriate connection, to obtain a considerable variety of input and output impedances.

3. *High gain.* Power gains of the order of 40 to 50 db per stage have been obtained

1

4. *Power handling capacity and efficiency.* The design can readily be varied to permit the required amount of power dissipation up to at least two watts. Furthermore the static characteristics are so nearly ideal that Class A efficiencies of 48 or 49 out of a possible 50% can be realized. The efficiencies for Class B and Class C operation are correspondingly high.

5. *Ruggedness and small size.* The germanium part of the transistor is enclosed in a hard plastic bead about $\frac{3}{16}$ inch in diameter. Inside the bead three connections are mechanically as well as electrically fastened to the germanium and are brought out as pigtails through the bead. This gives a very sturdy unit.

6. *Freedom from microphonics.* Vibration tests in the audio frequency range indicate that these devices are relatively free from microphonic noise.

7. *Limited frequency response.* Collector capacitance limits the frequency response at full gain to a few kilocycles. By using a suitable impedance
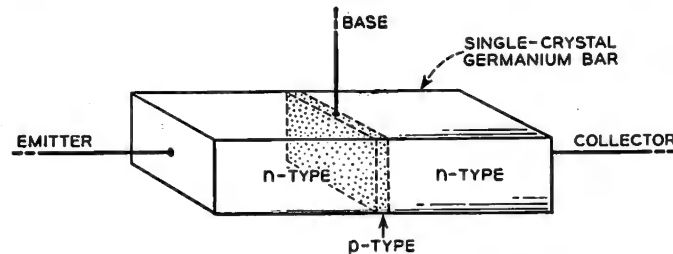
Fig. 1—The heart of an *n-p-n* transistor is a tiny bar of germanium to which three mechanically strong electrical connections are made.

mismatch it is possible to maintain the frequency response flat to at least one megacycle while still obtaining a useful amount of gain.

8. *Operation with exceedingly small power consumption.* Perhaps the most remarkable feature of these transistors is their ability to operate with exceedingly small power consumption. The best example of this to date is an audio oscillator which requires for a power supply only 6 microamperes at 0.1 volts. This represents 0.6 microwatts of power which contrasts sharply with the million or more microwatts required to heat the cathode of an ordinary receiving-type vacuum tube.

PHYSICAL APPEARANCE AND CONSTRUCTION

Figure 1 shows schematically the configuration of an *n-p-n* transistor. The small bar of single crystal germanium contains a thin layer of *p*-type interposed between regions of *n*-type. Mechanically strong ohmic connections are made to the three regions as indicated and brought out through a

2

hard plastic bead. A finished transistor is shown in the photograph of Fig. 2. It should be pointed out that Fig. 1 is not drawn to scale and that the p-layer may be less than a thousandth of an inch thick.

STATIC CHARACTERISTICS

A great deal of information about the low frequency performance of a transistor can be obtained from a set of static characteristics such as those shown in Fig. 4. Curves of this sort are obtained simply by connecting suit-
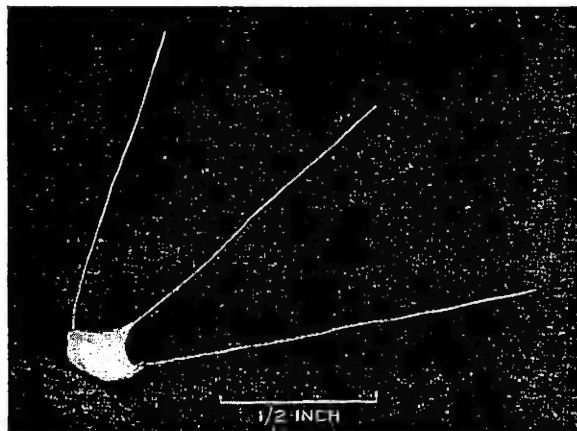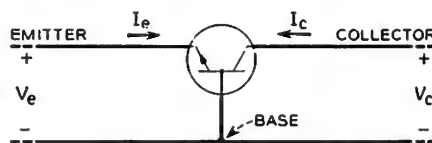

Fig. 2—A beaded n-p-n transistor.

Fig. 3—The symbol for a p-type transistor on which the convention of signs for currents and voltages is indicated.

able current sources to the emitter and collector circuits of the transistor and measuring the resulting voltages. The currents are called positive when they flow into the emitter and collector as shown and the voltages are called positive when they have the signs shown in Fig. 3.

Let us first examine these curves with an eye to finding out what kind of voltage and current supplies are needed to bias the transistor into the range in which it can amplify. To make this easy, that part of the characteristics which lies within the normal operating range has been shown as solid lines and that part of the characteristics corresponding to cutoff has been shown as dotted lines.

3

Note from the upper set of curves that V_c is positive in the operating range. This means that the collector must be biased positive with respect to the
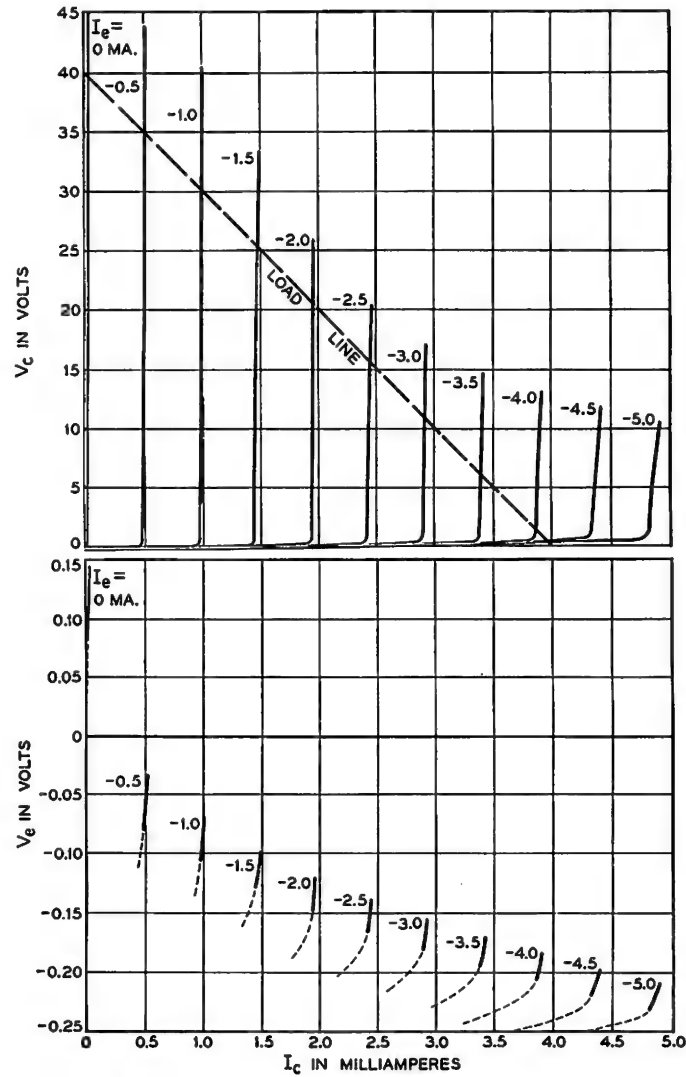

Fig. 4—Static characteristics of an *n-p-n* transistor.

base. For this particular transistor a bias voltage anywhere between about 0.1 volts and 35 volts is suitable. Note also that all the curves on this plot correspond to negative emitter currents. This means that the emitter must

4

be biased in such a way that current flows out of the emitter into a suitable current supply. Furthermore, the collector current corresponding to any given emitter current can be seen to be almost equal in magnitude to the emitter current. Since these two currents are opposite in sign, this means that most of the current which flows into the collector leaves by way of the emitter with the result that the current in the base circuit is very small.

Suppose that the collector is held at a constant positive voltage as, for example, by connecting a battery between collector and base (with a transformer winding in series, perhaps). Now if a negative current is forced into the emitter by a battery and resistance connected in series between emitter and base, the collector current can be controlled by varying the emitter current and will always be approximately equal in magnitude to the emitter current. Suitable collector currents for this particular transistor range from about 20 microamperes to about five milliamperes.

The exact choice of collector current and voltage within the ranges mentioned above will be dictated largely by the amount of power output required. The more power output required, the more current and voltage will be needed from the power supply. Since the collector circuit efficiency cannot exceed the theoretical limit of 50% in Class A operation, the signal power output cannot exceed half the power supplied by the battery. This means, for example, that if the collector is worked at 20 volts and 2 milliamperes the Class A power output cannot exceed 20 milliwatts.

From the lower plot of Fig. 4 it is possible to obtain information about the bias voltage required for the emitter. Note, first, that the entire emitter voltage plot corresponds to a very small range of emitter voltages near zero and, furthermore, that the part of the characteristics corresponding to the operating range covers only a few thousandths of a volt. This means that if the collector voltage is held constant very small changes in emitter voltage will produce fairly large changes in collector current, or if the collector current is held constant very small changes in emitter voltage will produce relatively enormous changes in collector voltage. This at once suggests the use of this transistor as a d-c. amplifier between a low impedance source and a high impedance load. In this application, voltage stepup of the order of 10,000 times is possible.

The very great sensitivity of the collector circuit to emitter voltage suggests, however, that for a-c. amplifiers one should use a current source as an emitter bias supply. This can be obtained from a battery and a large resistance in series. Furthermore, since the emitter voltage is always nearly zero, the emitter current can be calculated in advance by dividing the battery voltage by the value of the series resistance (provided, of course, that

5

the supply voltage is large compared to the few hundredths of a volt drop across the emitter circuit).

One can also draw some interesting conclusions from the static characteristics about the large signal operation of the transistor. If the load is resistive, the instantaneous operating point will swing up and down along a straight line such as the load line shown in the upper plot of Fig. 4. This particular load line corresponds to an a-c. load resistance of 10,000 ohms. Suppose that the steady collector biases are 20 volts and 2 milliamperes so that the drain from the power supply is 40 milliwatts. Now consider the permissible swings of collector voltage and current. Since the collector characteristics are quite straight and evenly spaced over a wide range of current and voltage values, the output signal can swing nearly down to zero collector volts and nearly up to zero collector current without distortion. The limit on the lower end is imposed by the fact that the collector characteristics begin to be curved when V_c is less than about 0.1 volts; and the limit on the upper end is imposed by the fact that the collector current does not drop completely to zero when I_e drops to zero. The lower limit of collector current is, in this case, about 50 microamperes and, since this amount of current in 10,000 ohms corresponds to 0.5 volts, this means that the instantaneous collector voltage is limited to swings between 39.5 volts and 0.1 volts. Starting from a quiescent value of 20 volts, the permissible positive swing is then 19.5 volts and the permissible negative swing is 19.9 volts. Reducing the quiescent voltage to 19.8 volts (and keeping the same load line) makes it possible to obtain a peak swing of 19.7 volts which corresponds to 19.45 milliwatts of signal delivered to the load. This gives a collector circuit efficiency of 48.5% out of a possible 50%. Some transistors take even less collector current when the emitter current is zero and hence permit even higher efficiencies.

These computations of efficiency have all been based on the assumption of sinusoidal *current* applied to the emitter. It will be shown in a later section that the emitter resistance varies with emitter current, however, and this means that to realize high efficiency with low distortion it is necessary to drive the emitter from a high impedance source.

OPERATION WITH SMALL POWER CONSUMPTION

For small signal applications the transistor represented by the characteristics of Fig. 4 can deliver useful gain at very much lower voltages and currents than those used in the example above. In order to show this, the characteristics of Fig. 5 have been plotted for a range of collector voltage extending up to only 2 volts and for a range of collector currents extending
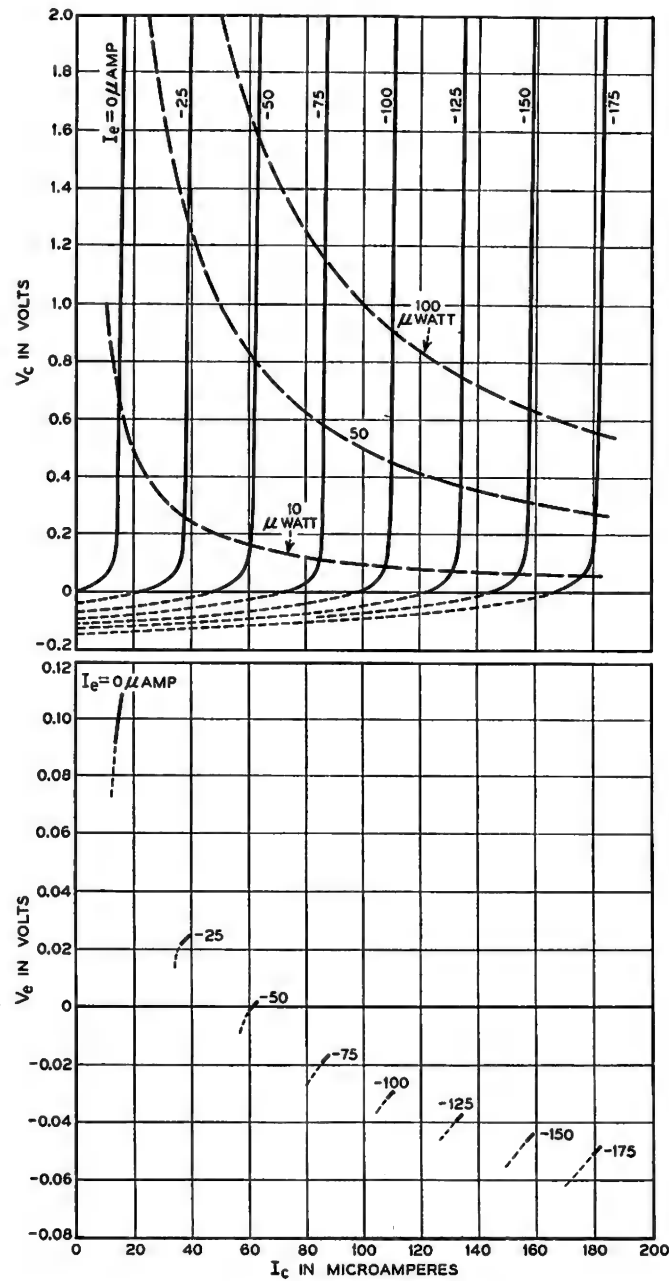
Fig. 5—Static characteristics showing behavior at very low applied voltages and currents.

7

up to only 200 microamperes. It can be seen from the upper plot that the collector circuit characteristics are still quite usefully straight and evenly spaced in this micro-power range. In fact, for small signal operation it is sufficient to use a collector voltage only a little in excess of 0.1 volts and a collector current a little in excess of 20 microamperes. This means that the power required to bias the collector into the operating range amounts to only a few microwatts. Contours are shown for 10, 50, and 100 microwatts of power supply.

This ability of the transistor to work with extremely small power consumption is one of its most striking and perhaps most important features.
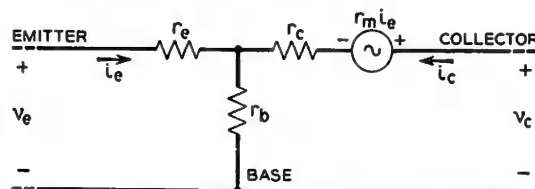
Fig. 6—The low-frequency equivalent circuit of a transistor.

When one considers that the total power consumption of a single transistor stage can be smaller by many thousands of times than the power required to heat the cathode in a vacuum tube, it is obvious that the advent of this device will make possible many new kinds of application.

VARIATION OF TRANSISTOR PROPERTIES WITH OPERATING POINT

Ryder and Kircher[5] have shown that it is convenient to analyze the small signal properties of a transistor at low frequencies in terms of the equivalent circuit of Fig. 6 where r_e is called the emitter resistance, r_b is called the base resistance, and r_c is called the collector resistance. The internal generator, $r_m i_e$, is the active part of the circuit and in this respect corresponds to the familiar μe_g of vacuum tube circuit theory. It is the purpose of this section to show what values these quantities have for a particular n-p-n transistor and to show how they vary with the applied biases. This will form a basis for the next section in which these quantities will be used to compute such things as the input and output impedances and the gains of various transistor connections.

Ryder and Kircher have shown that these four r's can be obtained directly from static characteristics such as those shown in Fig. 4 and Fig. 5. In the case of n-p-n transistors, however, the magnitudes of these quantities are such that it is difficult to obtain satisfactory accuracy in this way and it has been more convenient to measure the 4-pole r's by a-c. methods.

8

These measurements have shown that all of the r's are, to a first approximation, independent of collector voltage so long as the collector voltage is above a few tenths of a volt and so long as the total dissipation is small enough to prevent appreciable heating of the transistor.

In view of this fact it is perhaps sufficient to show how these quantities vary with emitter current for a moderate fixed value of collector voltage. Figures 7 and 8 show that r_c and r_m are very nearly equal and that they tend to decrease as I_e increases. Theoretically r_m and r_c should both be infinite. The fact that they reach values as low as 10 megohms in this case is a meas-
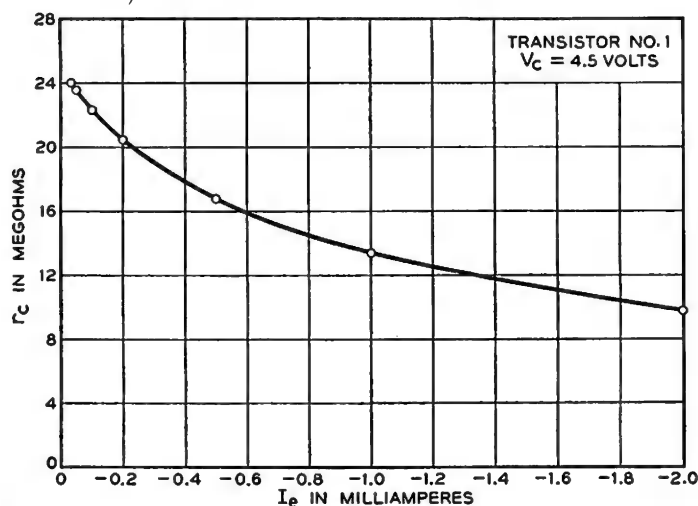
Fig. 7—The variation of collector resistance with emitter current at a fixed value of collector voltage.

ure of the imperfection in technique of fabricating the transistor. Values as high as 60 megohms have been achieved in the laboratory.

Figure 9 shows that r_b in this transistor is approximately 240 ohms and is independent of I_e.

Figure 10 shows that r_e decreases with increasing emitter current, ranging from about 500 ohms at 50 microamperes down to about 5 ohms at 5 milliamperes. Shockley[4] has shown that r_e should be given by

$$r_e = \frac{kT}{qI_e} \tag{1}$$

where q is the charge on an electron, k is Boltzman's constant, T is the Kelvin temperature and I_e is the emitter current. When the temperature
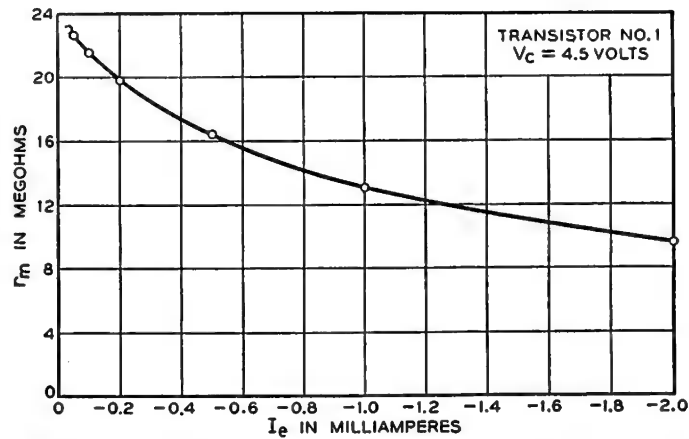
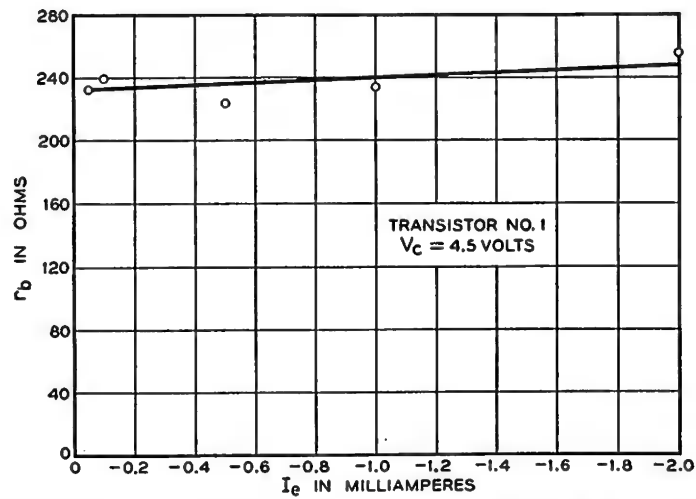Fig. 8—Variation of r_m with emitter current at fixed collector voltage.

Fig. 9—Variation of base resistance with emitter current. Scatter of the data indicates that the measurements were not accurate.

is about 80° F., this reduces to

$$r_e = \frac{25.9}{I_e} \qquad (2)$$

where I_e is measured in milliamperes. Within experiment error, values of r_e computed from this relation agree perfectly with the measured curve shown in Fig. 10.

Figure 11 introduces a new quantity, α, the current amplification factor of the transistor. This quantity is defined by the equation

$$\alpha = \frac{r_m + r_b}{r_c + r_b} \qquad (3)$$

Since r_m and r_c are both very large compared to r_b, α is approximately equal to the ratio of r_m to r_c. It will be shown in a later section that this quantity is important in determining some of the circuit properties of the transistor and that many of the circuit properties become more desirable as α approaches unity.
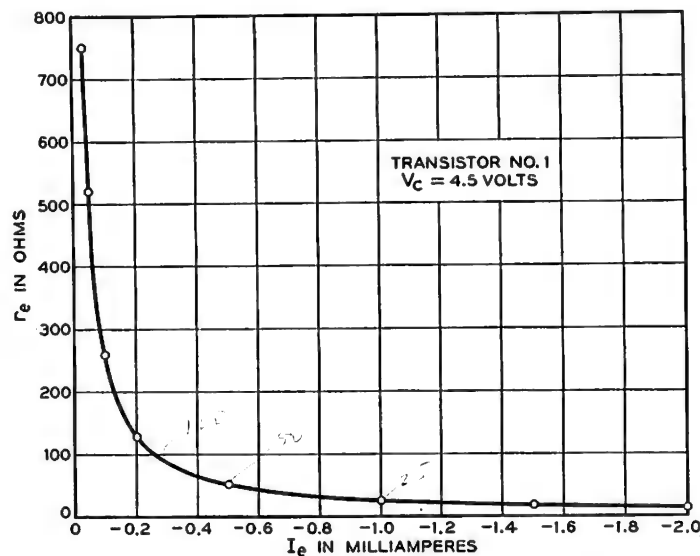
Fig. 10—The emitter resistance is inversely proportional to emitter current.

It can be seen from Fig. 11 that in this transistor α is approximately equal to 0.98 and that it increases slightly with increasing emitter current. The highest value of α so far achieved is 0.9965.

Those units which have been made in the laboratory so far show considerable variation in some of the properties, but this is partly due to the fact that changes have been made deliberately to test one aspect or another of Shockley's theory. The data in table I are presented to indicate what properties have been achieved to date. The collector capacitance C_c will be discussed in a later section.

General Considerations and Formulae

It is a consequence of the fact that α is always less than unity in this ·structure that these transistors are unconditionally stable with all termina-

tions. This means that stability considerations do not prevent working with matched terminations. Furthermore it is possible to obtain a variety of input and output impedances by connecting the transistor as a grounded-emitter, grounded-base, or grounded-collector stage. It is the purpose of this section to give some idea of the characteristics of these various stages and to show in each case at least one way of supplying the required biases and couplings to the stage.
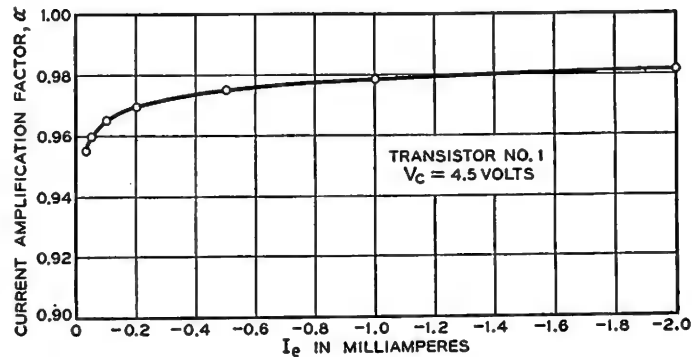

Fig. 11—The current amplification factor, α, increases slightly with increasing emitter current. Note the expanded scale for α.

TABLE I

CONSTANTS FOR VARIOUS TRANSISTORS MEASURED AT $V_c = 4.5$ v., $I_e = 1.0$ ma.

Transistor No.		I	II	III	IV	V
r_e	(ohms)	25.9	31.6	33.1	30.2	38.8
r_b	(ohms)	240	44	300	3070	180
r_c	(megohms)	13.4	0.626	1.11	1.21	2.00
$r_c - r_m$	(megohms)	0.288	0.00387	0.0168	0.00422	0.0439
α		0.9785	0.9936	0.9848	0.9965	0.9780
C_c	($\mu\mu$f.)	7	7.7	18.9	27.9	21.2

It will be convenient to begin by writing down general relationships which will apply to all the possible connections. To this end let the transistor be represented by the box in Fig. 12. At low frequencies, the signal currents and voltages are related through the equations:

$$R_{11}i_1 + R_{12}i_2 = v_1$$
$$R_{21}i_1 + R_{22}i_2 = v_2 \tag{4}$$

If a generator of open circuit voltage v_g and internal resistance R_g is connected to the input terminals of the device as shown in Fig. 13, then

$$v_1 = v_g - i_1R_g \tag{5}$$

12

and if a load of resistance R_L is connected to the output terminals

$$v_2 = -R_L i_2 . \tag{6}$$

The equations for the circuit of Fig. 13 are, therefore,

$$(R_{11} + R_g)i_1 + R_{12}i_2 = v_g$$
$$R_{21}i_1 + (R_{22} + R_L)i_2 = 0 \tag{7}$$

Solving for the voltage developed across the load $(= -R_L i_2)$ gives

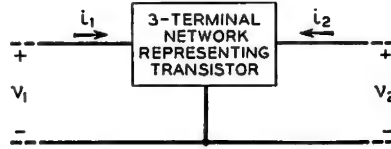$$v_2 = \frac{R_L R_{21}}{(R_{11} + R_g)(R_{22} + R_L) - R_{12}R_{21}} v_g \tag{8}$$

Fig. 12—A three-terminal network representing either grounded emitter, grounded base, or grounded collector connection of a transistor. Note the convention of signs.
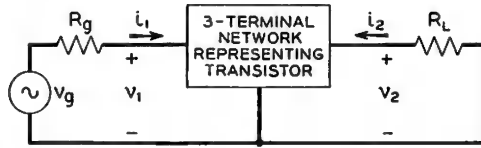
Fig. 13—The three-terminal network of Fig. 12 connected between a generator and a load.

The power gain in the circuit is the power delivered to the load (v_2^2/R_L) divided by the power available from the generator $(v_g^2/4R_g)$. From equation (8), this gives

$$G = \frac{4R_g R_L R_{21}^2}{[(R_{11} + R_g)(R_{22} + R_L) - R_{12}R_{21}]^2} \tag{9}$$

The gain depends on R_g and R_L and will be maximum when these are chosen to match the input and output impedances of the transistor stage. But the input impedance depends on R_L and the output impedance depends on R_g in the following way:

$$\text{Input impedance} = R_i = R_{11} - \frac{R_{12}R_{21}}{R_{22} + R_L} \text{ and} \tag{10}$$

$$\text{Output impedance} = R_o = R_{22} - \frac{R_{12}R_{21}}{R_{11} + R_g} \tag{11}$$

13

If $R_i = R_g$ and $R_o = R_L$ then impedances are matched at the input and output terminals and the gain is a maximum. The conditions are:

Matched input impedance =

$$R_{im} = R_{11} \sqrt{1 - R_{12} R_{21}/R_{11} R_{22}}, \tag{12}$$

Matched output impedance =

$$R_{om} = R_{22} \sqrt{1 - R_{12} R_{21}/R_{11} R_{22}}, \tag{13}$$

Maximum available gain =

$$\text{M.A.G.} = \frac{R_{21}^2}{R_{11} R_{22}} \frac{1}{[1 + \sqrt{1 - R_{12} R_{21}/R_{11} R_{22}}]^2} \tag{14}$$

THE GROUNDED BASE STAGE

In this and the following two sections we will put into equations (7) through (14) the appropriate 4-pole r's to obtain expressions for impedances and gains. As a numerical example we will substitute into the resulting equations the measured values of these r's for Transistor No. I working at $V_c = 4.5$v. and $I_c = 1$ ma. It must be understood that the numerical values may vary appreciably from transistor to transistor and that these numerical calculations are intended only for illustration and not as a basis for final circuit design. The numerical values to be used are

$$r_e = 25.9 \text{ ohms}$$

$$r_b = 240 \text{ ohms}$$

$$r_c = 13.4 \ (10)^6 \text{ ohms} \tag{15}$$

$$r_c - r_m = 0.288 \ (10)^6 \text{ ohms}$$

$$\alpha = 0.9785$$

In this section it will be shown that the grounded base connection is suitable for working between a low impedance source and a high impedance load. The input impedance may be of the order of a hundred ohms and the output impedance of the order of one or more megohms. In this connection the current amplification is always less than unity but the voltage amplification may be very large indeed. Power gains of the order of 40 to 50 db can be obtained between matched impedances, and appreciable gains can still be obtained if the load resistance is reduced to a few thousand ohms (because the current gain is then almost equal to unity). In this case the gain of the stage is almost completely independent of those transistor properties

14

which tend to vary from unit to unit. This sort of stage does not produce a phase reversal.
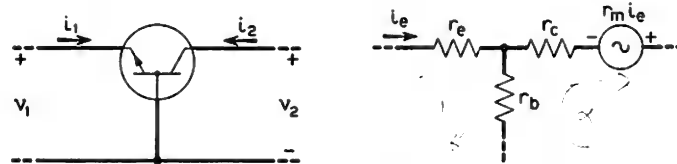

Fig. 14—The grounded base connection of a transistor.

For the grounded base stage shown in Fig. 14,

$$R_{11} = r_e + r_b = 266 \text{ ohms}$$

$$R_{12} = r_b = 240 \text{ ohms}$$

$$R_{21} = r_m + r_b = 13.1 \ (10)^6 \text{ ohms}$$

$$R_{22} = r_c + r_b = 13.4 \ (10)^6 \text{ ohms} \tag{16}$$

$$\alpha = \frac{r_m + r_b}{r_c + r_b}$$

$$\doteq \frac{r_m}{r_c} \approx \alpha$$

$$= 0.9785$$

In this case if r_b is neglected by comparison with r_m and r_c, equation (8) leads to

$$v_2 = \frac{\alpha R_L v_g}{(r_e + r_b + R_g)(1 + R_L/r_c) - \alpha r_b}. \tag{17}$$

Since for these transistors $\dfrac{r_m}{r_c}$ ($\doteq \alpha$) is always less than unity, the output voltage is in phase with the input voltage. Furthermore, if R_L is very high, the output voltage is enormous by comparison with the input voltage. For example, if $R_g = 0$ and R_L is infinite

$$v_2 = v_g \frac{r_m}{r_e + r_b} \tag{18}$$

and for the numerical example this is

$$v_2 = 4.93 \ (10)^4 \ v_g.$$

15

To achieve this step-up would require a load impedance very large compared to 13 megohms, but even with more modest values of load impedance the voltage step-up is large.

If R_L is small compared to r_c, the second of equations (7) leads to

$$i_2 = -\frac{r_m}{r_c} i_1$$

$$\doteq -i_1 \tag{19}$$

[handwritten: $\propto \dfrac{r_c}{r_c + R_L}$]

and the current delivered to the load is approximately equal to the current which the generator delivers to the transistor.

From equations (10) and (11), the input and output impedances are

$$R_i = r_e + r_b - \frac{r_b(r_m + r_b)}{r_c + R_L + r_b} \tag{20}$$

[handwritten: $\doteq r_b \dfrac{(1-\alpha) + \frac{R_L}{r_c}}{1 + R_L/r_c} + r_e$]

$$R_o = r_c + r_b - \frac{r_b(r_m + r_b)}{r_e + r_b + R_g} \tag{21}$$

[handwritten: $= r_c \left(1 + \dfrac{r_b \left(\frac{R_g}{r_c} - \alpha \right)}{r_e + r_b + R_g} \right)$]

As the load impedance varies from zero to infinity, the input impedance varies from

$$R_i = r_e + r_b \left[1 - \frac{r_m + r_b}{r_c + r_b} \right] \quad \text{for } R_L = 0$$

$$\doteq r_e + r_b(1 - \alpha)$$

$$= 31.1 \text{ ohm}$$

[handwritten: (22) R_L ...]

to

$$R_i = r_e + r_b = 266 \text{ ohms} \quad \text{for } R_L = \infty. \tag{23}$$

When $R_g = 0$, the output impedance is

$$R_o = r_c - \frac{r_b}{r_e + r_b}(r_m - r_e) \tag{24}$$

[handwritten: This has all the impedance as transistor and upon it.]

$$\doteq r_c - \frac{r_b}{r_e + r_b} r_m = r_c \left(1 - \frac{\alpha\, r_b}{r_e + r_b} \right) \tag{25}$$

[handwritten: and upon it.]

$$= 1.56 \,(10)^6 \text{ ohms.}$$

As R_g increases to infinity

$$R_o = r_c + r_b = 13.4 \,(10)^6 \text{ ohms.} \tag{26}$$

[handwritten bottom: For Equations also see Electr. Eng. Short March 1953 p. ...]

From equations (12) and (13), the matched input and output impedances are approximately

$$R_{im} = (r_e + r_b) \sqrt{1 - \alpha r_b/(r_e + r_b)}$$
$$= 91 \text{ ohms} \tag{27}$$

$$R_{om} = (r_c + r_b) \sqrt{1 - \alpha r_b/(r_e + r_b)}$$
$$= 4.58(10)^6 \text{ ohms.} \tag{28}$$

With matched impedances, the maximum available gain is

$$\text{M.A.G.} = \frac{\alpha(r_m + r_b)}{r_e + r_b} [1 + \sqrt{1 - \alpha r_b/(r_e + r_b)}]^{-2}$$
$$= 2.7 \ (10)^4 \text{ or } 44.3 \text{ db.} \tag{29}$$

The matched output impedance of this stage is inconveniently high but a useful amount of gain can be maintained if R_L is reduced to a more rea-
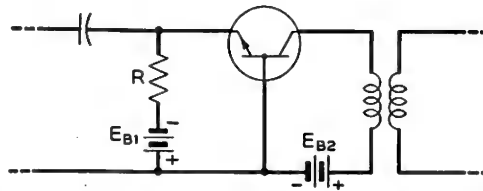
Fig. 15—One practical arrangement of a grounded base amplifier stage.

sonable value. For example, if $R_L = 200,000$ and $R_g = 25$, equation (9) gives

$$G = 5.3 \ (10)^3 \text{ or } 37.2 \text{ db.}$$

If stages of this sort are to be cascaded, a step-down transformer must be used to couple each collector to the following emitter. Otherwise, since the current amplification factor of the transistor is slightly less than unity, the gain per stage will also be slightly less than unity.

One practical arrangement of a grounded base stage would be as shown in Fig. 15. The required value of R will be approximately

$$R = \frac{E_{B1}}{I_c} \tag{30}$$

where I_c is the desired collector current and E_{B1} is the voltage of the emitter-bias battery. For operating at $I_c = 1$ ma, for example, $E_{B1} = 1.5$ v and $R = 1500$ ohms would be suitable.

17

For many applications the grounded emitter connection is more desirable than either of the other two. The power gains which can be obtained are high—of the order of 50 db—and the interstage coupling problem is simplified by the fact that the input impedance is somewhat higher than that of the grounded base stage while the output impedance is very much lower. The input impedance may be of the order of a few hundred ohms and the output impedance of the order of a few hundred thousand ohms. Both voltage and current amplification are produced (with a phase reversal) and gains of the order of 30 db or more per stage can be obtained without the
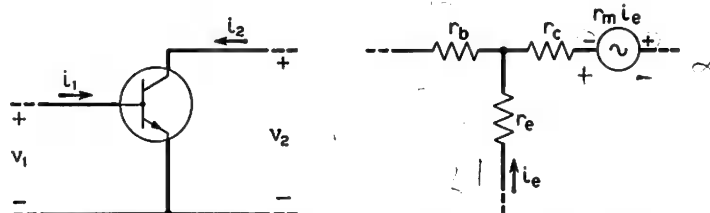
Fig. 16—The grounded emitter connection of a transistor and the equivalent circuit

use of interstage coupling transformers. The input and output impedances depend very critically on α and may vary appreciably from unit to unit.

For this connection, which is indicated schematically in Fig. 16,

$$R_{11} = r_e + r_b = 266 \text{ ohms}$$

$$R_{12} = r_e = 25.9 \text{ ohms}$$

$$R_{21} = r_e - r_m = -13.1 \ (10)^6 \text{ ohms} \tag{31}$$

$$R_{22} = r_e + r_c - r_m = 0.288 \ (10)^6 \text{ ohms}$$

Putting these values into equation (8) shows v_2 is always opposite in sign compared with v_g, that is, that the grounded emitter stage produces a phase reversal as does the grounded cathode vacuum tube.

If R_L is infinite and $R_g = 0$

$$v_2 = v_g \frac{r_e - r_m}{r_e + r_b} \quad \approx v_g \frac{-\alpha \, r_c}{r_e + r_b}$$

$$= -4.93 \ (10)^4 \, v_g$$

18

which is the same as for the grounded base stage. But if $R_L = 0$

$$i_2 = \frac{r_m - r_e}{r_e + r_c - r_m} i_1 \qquad (32)$$

$$\doteq \frac{\alpha}{1 - \alpha} i_1 \qquad (33)$$

$$= 45.5\, i_1.$$

Thus it is seen that the grounded emitter amplifier can produce quite appreciable current amplification—particularly so when α approaches unity.

The input impedance to the stage is

$$R_i = r_e + r_b + \frac{r_c(r_m - r_e)}{r_e + r_c - r_m + R_L}. \qquad (34)$$

$$r_b + r_e\, \frac{r_c}{r_c(1-\alpha) + R_L}$$

When $R_L = 0$ this reduces to

$$R_i = r_b + r_e \frac{1}{\dfrac{r_e}{r_c} + 1 - \dfrac{r_m}{r_c}} \qquad (35)$$

$$\doteq r_b + r_e \frac{1}{1 - \alpha} \qquad (36)$$

$$= 1440 \text{ ohms.}$$

As R_L increases to infinity, the input impedance decreases to $r_e + r_b$ which, for the numerical example, is 266 ohms.

The output impedance is

$$R_o = r_e + r_c - r_m + \frac{r_e(r_m - r_e)}{r_e + r_b + R_g} \qquad (37)$$

When $R_g = 0$, this gives

$$R_o = r_c - \frac{r_b}{r_e + r_b}(r_m - r_e) \qquad (38)$$

$$r_c\left(1 - \alpha\, \frac{r_b + R_g}{r_e + r_b + R_g}\right)$$

$$\doteq r_c\left[1 - \frac{r_b}{r_e + r_b}\alpha\right] \qquad (39)$$

$$= 1.56\,(10)^6 \text{ ohms}$$

As R_g increases to infinity, R_o decreases to

$$R_o = r_e + r_c - r_m \qquad (40)$$

$$= 0.288\,(10)^6 \text{ ohms.}$$

19

The matched input and output impedances are

$$R_{im} = (r_e + r_b) \sqrt{1 + r_e(r_m - r_e)/(r_e + r_b)(r_e + r_c - r_m)} \qquad (41)$$

$$= 619 \text{ ohms and}$$

$$R_{om} = (r_e + r_c - r_m) \sqrt{1 + r_e(r_m - r_e)/(r_e + r_b)(r_e + r_c - r_m)} \qquad (42)$$

$$= 0.671 \,(10)^6 \text{ ohms}$$

As α increases toward unity the matched input impedance increases and the matched output impedance decreases. They approach the limits

$$R_{im} = \sqrt{(r_b + r_c)(r_e + r_b)} \qquad (43)$$

$$R_{om} = r_c \sqrt{(r_b + r_c)/(r_e + r_b)} \qquad (44)$$

$$\text{as } \alpha \to 1.$$

If r_m in the transistor of our numerical examples could be increased to exactly the value of r_c ($\alpha = 1$) then the matched impedances would be

$$R_{im} = 59,700 \text{ ohms}$$

$$R_{om} = 5,800 \text{ ohms}$$

From this example, it is seen that the impedances vary rapidly with α as α approaches unity.

With matched impedances, the maximum available gain from the grounded emitter stage is

$$\text{M.A.G.} = \frac{(r_c - r_m)^2}{r_e r_c} \left[\sqrt{\left(1 + \frac{r_b}{r_e}\right)\left(\frac{r_e}{r_c} + 1 - \frac{r_m}{r_c}\right)} + \sqrt{1 + \frac{r_b}{r_e}\left(\frac{r_e}{r_c} + 1 - \frac{r_m}{r_c}\right)} \right]^{-2} \qquad (45)$$

$$= 2.02 \,(10)^5 \text{ or } 53 \text{ db}$$

When α is exactly unity this expression reduces to r_c/r_e provided r_e and r_b are small compared with r_c. For values of α which are enough smaller than unity so that

$$\frac{r_e}{r_c} \ll 1 - \alpha$$

the expression for maximum available gain reduces to the approximate expression

$$\text{M.A.G.} = \alpha(r_m/r_e) \left[\sqrt{(1 - \alpha)\frac{r_b}{r_e}} + \sqrt{1 + (1 - \alpha)\frac{r_b}{r_e}} \right]^{-2} \qquad (46)$$

From the above equations it can be seen that gain does not increase rapidly with α when α is sufficiently near unity. In the case of our numerical example, increasing α from 0.9785 to unity increases the gain by only 4.1 db. The gain of the stage is approximately proportional to r_c and inversely proportional to r_e and can therefore be increased by operating at higher emitter currents or by fabricating the transistor in such a way as to obtain higher values of r_c . In the latter case it would be desirable also to increase α in order to keep the output impedance from becoming unreasonably high.

It has been seen that, for the case of our numerical example, the matched output impedance is large compared to the input impedance (671,000 ohms compared to 619 ohms). This means that if the maximum available gain is to be obtained in cascaded stages, step-down interstage transformers must be used. But an appreciable amount of gain can be obtained without interstage impedance matching. This is because of the short-circuit current amplification previously mentioned which amounts approximately to

$$\frac{\alpha}{1 - \alpha}$$

or 45.5 times in the case of our numerical example. For this transistor, then, the iterative gain per stage without impedance transformation would be 33.2 db. This gain increases very rapidly as α approaches unity, not only because the short-circuit current amplification increases, but also because the output impedance decreases and the input impedance increases so that a better interstage impedance match is obtained. From equations (41) and (42), it is seen that the matched input and output impedances are equal when

$$r_c - r_m = r_b$$

or when

$$1 - \alpha = \frac{r_b}{r_c + r_b}$$

In this case the gain per stage would be approximately r_c/r_e . This says that if r_m could be increased in the numerical example until

$$\alpha = 0.999821$$

the gain per stage (without impedance transformation) would be increased to 57.1 db. Values of α this near to unity have not even been approached in transistors made to date. This unrealistic example is included only to indicate one of the reasons for seeking to make α very near unity.

Consider next one possible way of supplying biases to a grounded emitter

21

stage. Suppose that the collector is connected through a transformer winding to a fixed voltage supply as indicated in Fig. 17. Since V_e is always a very small fraction of a volt (see Fig. 4 and Fig. 5) the collector voltage will be approximately equal to the supply voltage. If no d-c. connection is supplied to the base, it will float at a potential above ground equal (in magnitude) to V_e—i.e., a very small fraction of a volt—and the collector current will be exactly equal to the emitter current. To find out approximately what this value of the current will be, consider the upper set of static characteristics in Fig. 5. Note that, when the emitter current is zero, the collector current is of the order of 20 microamperes (the exact value varies from 1 to 30 microamperes in transistors tested so far). The collector current is then of the order of 20 microamperes greater than the emitter current. If the emitter current is now increased by ΔI_e, the collector current will increase by $\alpha \Delta I_e$. That is, the increments of collector current will be slightly smaller than the incre-
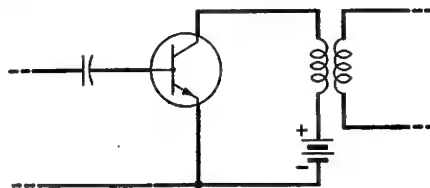

Fig. 17—One practical arrangement of a grounded emitter stage.

ments of emitter current and, as the emitter current is increased, the emitter and collector currents will become more nearly equal. If α were perfectly constant they would become exactly equal when

$$I_e = I_c = \frac{I_{co}}{1 - \alpha} \tag{47}$$

where I_{co} is the collector current which flows when the emitter current is zero. Equation (47), then, gives the value of emitter (and collector) current which will flow in a grounded emitter stage if no d-c. connection is made to the base. This current varies rapidly with α. For the transistor which has been considered numerically, this current would amount to about 465 microamperes which is certainly within the range of suitable values for the transistor. In certain low level applications, however, it might be desirable to work at a smaller current for the sake of decreasing battery power consumption. This requires that a small current be drawn out of the base. The required current is small because the collector current will decrease by $1/(1 - \alpha)$ microamperes for each microampere drawn from the base. One method of obtaining this base current is to provide a resistive path between

22

base and ground as shown, for example, in Fig. 18. Since the base floats at a positive potential with respect to ground, this circuit produces a base current of the right sign to decrease the collector current. As the value of the series resistance is decreased to zero, the collector current decreases to
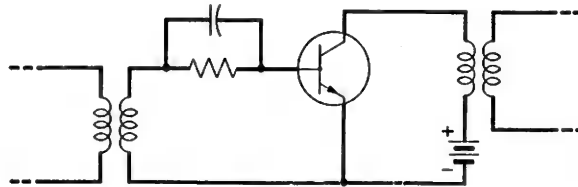

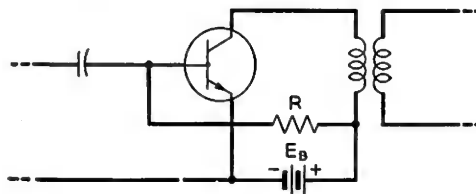Fig. 18—Modification of Fig. 17 to obtain lower collector current.

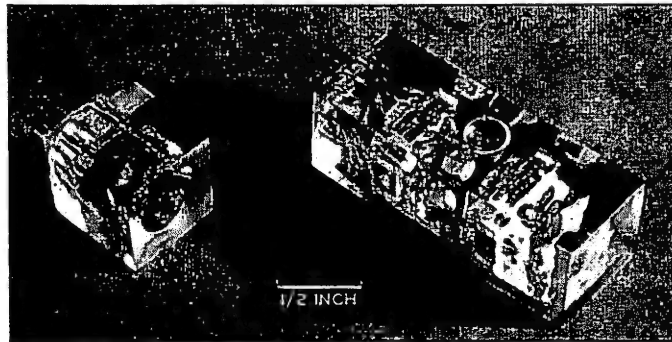Fig. 19—Modification of Fig. 17 to obtain higher collector current.

Fig. 20—A two-stage grounded emitter amplifier which produces approximately 90 db power gain is shown on the right and a micro-power audio oscillator is shown on the left.

a value corresponding to zero emitter voltage. A still further decrease in collector current can be obtained by inserting resistance between emitter and ground.

In order to increase the collector current to values higher than that corresponding to zero base current, a high resistance path between base and the positive supply voltage may be used as shown in Fig. 19. In this case the

collector current will increase by $1/(1 - \alpha)$ microamperes for each microampere which flows through the bias resistor. Since the current in the bias resistor will be approximately E_B/R, it is a simple matter to compute the required value of bias resistor once the desired collector current is known.

Figure 20 shows a two-stage audio amplifier which gives approximately 90 db gain. The circuit is shown in Fig. 21.
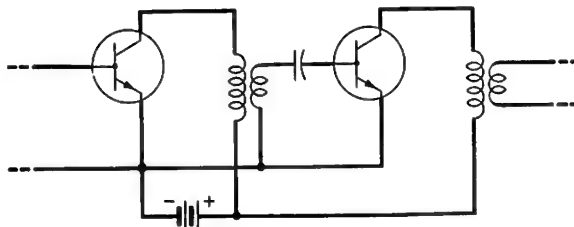

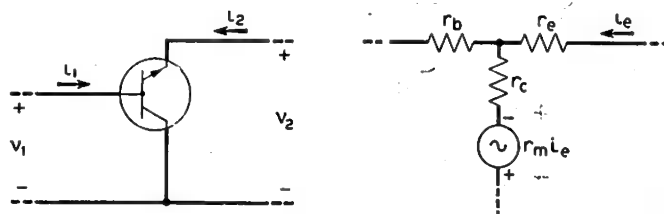Fig. 21—Circuit of the amplifier shown in Fig. 20.

Fig. 22—The grounded collector connection of a transistor and the equivalent circuit.

THE GROUNDED-COLLECTOR STAGE

Although the power gain obtainable from this connection is relatively low—of the order of 15 or 20 db—it has very interesting possibilities in producing very high input impedances or very low output impedances. If it is worked into a fairly high load impedance, the input impedance may be several megohms, or if it is worked from a source of moderately low impedance (a few thousand ohms), the output impedance may be of the order of 25 ohms or lower.

For this type of stage, which is shown schematically in Fig. 22,

$$R_{11} = r_b + r_c = 13.4 \ (10)^6 \ \text{ohms}$$

$$R_{12} = r_c - r_m = 0.288 \ (10)^6 \ \text{ohms}$$

$$R_{21} = r_c = 13.4 \ (10)^6 \ \text{ohms} \tag{48}$$

$$R_{22} = r_e + r_c - r_m = 0.288 \ (10)^6 \ \text{ohms}$$

24

If this stage is worked from a zero impedance generator into an infinite impedance load

$$v_2 = v_g \left(\frac{r_c}{r_b + r_c} \right) \tag{49}$$

$$\doteq v_g$$

and so, like a cathode follower, it gives an output voltage which is less than the input voltage, but in the same phase.

If the stage is operated into a short circuit

$$i_2 = -i_1 \frac{r_c}{r_e + r_c - r_m} \tag{50}$$

$$\frac{i_2}{i_1} \doteq \frac{-1}{1 - \alpha + R_L/r_c}$$

$$\doteq -i_1 \frac{1}{1 - \alpha} \tag{51}$$

$$= -46.5 \, i_1$$

which indicates that the stage can give an appreciable current gain.

The input impedance is

$$R_i = r_b + r_e - \frac{r_c(r_c - r_m)}{r_e + r_c - r_m + R_L} \tag{52}$$

$$\doteq r_b + \frac{r_e + R_L}{(1 - \alpha) + R_L/r_c}$$

When $R_L = 0$, this reduces to

$$R_i = r_b + r_e \frac{1}{\frac{r_e}{r_c} + 1 - \frac{r_m}{r_c}} \tag{53}$$

$$\doteq r_b + r_e \frac{1}{1 - \alpha} \tag{54}$$

$$= 1445 \text{ ohms.}$$

When R_L is infinite

$$R_i = r_b + r_c$$
$$= 13.4(10)^6 \text{ ohms.} \tag{55}$$

With respect to input impedance, the grounded collector stage is again seen to be like a cathode follower in that the input impedance is high when the load impedance is high.

The output impedance is

$$R_o = r_e + r_c - r_m - \frac{r_c(r_c - r_m)}{r_b + r_e + R_g} \tag{56}$$

25

For $R_g = 0$, this reduces to

$$R_o = r_e + r_b \frac{r_c - r_m}{r_b + r_c} \doteq r_\varepsilon + (1 - \alpha) \frac{r_b + R_g}{1 + R_g/r_c} \quad (57)$$

$$\doteq r_e + r_b (1 - \alpha)$$

$$= 31.1 \text{ ohms.} \quad (58)$$

For R_g infinite

$$R_o = r_e + r_c - r_m$$

$$= 0.288(10)^6 \text{ ohms.} \quad (59)$$

The matched input impedance is

$$R_{im} = (r_b + r_c) \sqrt{\frac{r_b}{r_b + r_c} + \frac{r_e r_c}{(r_b + r_c)(r_e + r_c - r_m)}} \quad (60)$$

$$\doteq \sqrt{r_c[r_b + r_e/(1 - \alpha)]}$$

$$= 139,000 \text{ ohms.} \quad (61)$$

The matched output impedance is

$$R_{om} = (r_e + r_c - r_m) \sqrt{\frac{r_b}{r_b + r_c} + \frac{r_e r_c}{(r_b + r_c)(r_e + r_c - r_m)}} \quad (62)$$

$$\doteq (1 - \alpha) \sqrt{r_c[r_b + r_e/(1 - \alpha)]}$$

$$= 2990 \text{ ohms.} \quad (63)$$

With matched impedance, the maximum available gain of the grounded collector stage is

$$\text{M.A.G.} = \frac{\dfrac{r_c^2}{(r_b + r_c)(r_e + r_c - r_m)}}{\left[1 + \sqrt{\dfrac{r_b}{r_b + r_c} + \dfrac{r_e r_c}{(r_b + r_c)(r_e + r_c - r_m)}}\right]^2} \quad (64)$$

As α approaches unity, this approaches approximately

$$\text{M.A.G.} = r_c/4r_e \quad (65)$$

but so long as $r_e \ll r_c - r_m$, a good approximation is

$$\text{M.A.G.} = 1/(1 - \alpha)$$

$$= 46.5 \text{ or } 16.7 \text{ db.} \quad (66)$$

26

The considerations involved in supplying biases to a grounded collector stage are rather similar to those discussed already for the grounded emitter case. If the base is allowed to float, the collector current will be given approximately by equation (47) as discussed for the grounded emitter case. A resistance between base and the negative side of the supply battery in Fig. 24 will serve to decrease the collector current while a resistance between base and ground will serve to increase it. In applications where it is desired to make full use of the high input impedance which this stage can afford, it may be most desirable to let the base float as shown in Fig. 23.
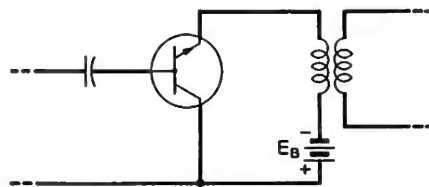
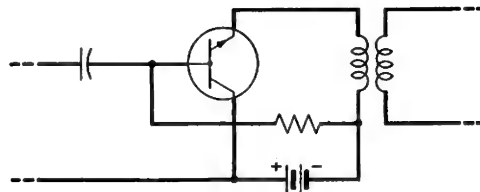Fig. 23—One practical arrangement of a grounded collector stage.

Fig. 24—Modification of Fig. 23 to obtain lower collector current. To raise collector current remove the resistance shown and connect a high resistance between base and ground.

FREQUENCY RESPONSE—GENERAL REMARKS

Shockley has shown that there are several different physical considerations which lead one to expect a high-frequency cutoff in the response of n-p-n transistors. The frequency at which cutoff occurs depends in a theoretically understandable way on such things as the geometry of the transistor and the physical properties of the germanium from which it is made. If these factors could all be controlled and varied at will, it would be possible to design a transistor to have a specified cutoff frequency.

One limitation comes about in the following way: In order to produce transistor action, the electrons which are injected into the p layer at the emitter junction must travel across this thin layer and arrive at the collector junction. They do this principally by a process of diffusion and require a finite (but small) amount of time to make the journey. If this time were

exactly the same for all electrons, the effect would be simply to delay the output signal with respect to the input and there would be no effect on frequency response. But there is a certain amount of dispersion in transit time which means that the electrons corresponding to a particlar part of the input signal wave do not all arrive simultaneously at the collector. When this difference in time of arrival amounts to an appreciable part of a cycle there is a tendency for some of the electrons to cancel the effect of others so that the frequency response begins to fall off. As the signal frequency increases beyond this point, the effect becomes more and more pronounced and the response continues to fall with increasing frequency.

In terms of the equivalent circuit, this dispersion in transit time means that beyond a certain frequency, r_m (and hence α) begins to decrease with increasing frequency and so the transistor may be said to have a certain α-cutoff which we will call $f_{c\alpha}$.

Shockley has shown that $f_{c\alpha}$ is inversely proportional to the square of the p-layer thickness and hence increases rapidly as the p layer is made thinner. For n-p-n transistors now available, this cutoff should occur at frequencies between five and twenty megacycles.

Another limitation on frequency response comes about from the fact that, at sufficiently high frequencies, the emitter junction fails to behave as a pure resistance and is, in effect, shunted by a capacitance. In terms of the equivalent circuit, this means that r_e is shunted by a capacitance.

The effect which this has on frequency response can be reduced by reducing the impedance of the source from which the emitter is driven. But so far as the emitter junction is concerned, r_b is always in series with the source impedance and so it is the value of r_b which ultimately determines the emitter cutoff frequency.

This capacitative reactance should begin to become appreciable with respect to emitter resistance at a frequency which may be of the same order as $f_{c\alpha}$. If r_b is high, the emitter cutoff frequency f_{ce} will then be of the same order of magnitude as $f_{c\alpha}$ and will increase as r_b is decreased.

A third cause for limited frequency response is the capacitance of the collector junction. The n-type germanium on one side of the junction behaves as one plate of a parallel-plate condenser and the p-type germanium on the other side behaves as the other plate. Since the transition from n to p type germanium may be made in an exceedingly small fraction of an inch, the plates of the condenser are very closely spaced and the capacitance may be appreciable.

Collector capacitance also depends on collector voltage, decreasing with increasing voltage. Theoretically, the capacitance should be in proportion to the negative one-third power of V_c:
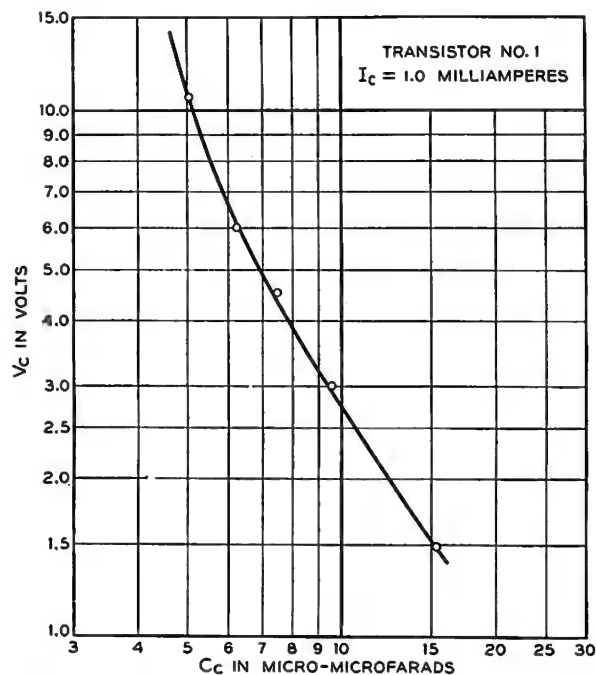
28

Fig. 25—Collector capacitance decreases as collector voltage is increased.
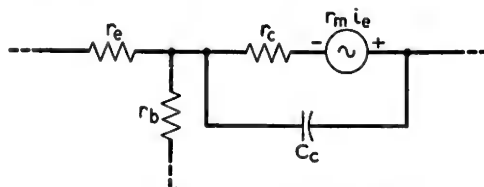

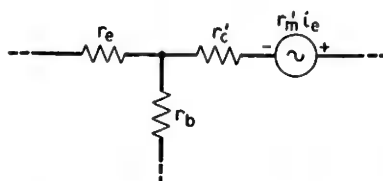Fig. 26—The equivalent circuit of a transistor with collector capacitance shown.

Fig. 27—The effect of collector capacitance is to change r_m and r_c to r_m' and r_c'. See equations (67) and (68).

Figure 25 shows measured values of C_c as a function of collector voltage. For reasons which are not understood at present, these data show a departure from the usual inverse one-third power variation. At $V_c = 4.5$ volts

the capacitance is seen to be approximately 7 micro-microfarads. In terms of the equivalent circuit, this capacitance is in shunt with the series combination of r_c and the generator, $r_m i_e$, as shown in Fig. 26. This can be shown to be equivalent to the circuit of Fig. 27 in which r_c has been replaced by

$$r_c' = r_c/(1 + jC_c r_c \omega) \tag{67}$$

and r_m has been replaced by

$$r_m' = r_m/(1 + jC_c r_c \omega). \tag{68}$$

The effect of collector capacitance can be computed by substituting r_c' and r_m' for the values of r_m and r_o (implicitly contained) in equation (8). In the sections which follow, this will be done for each of the three transistor connections and the resulting collector cutoff frequencies f_{cc} will be computed. It will be shown that at least for the transistor on which data are presented collector capacitance tends to produce a cutoff frequency well below those to be expected from emitter cutoff or alpha cutoff. For this reason, only collector cutoff will be considered.

COLLECTOR CUTOFF IN THE GROUNDED BASE STAGE 1.2×10^8

If the values of r_c' and r_m' from equations (67) and (68) are substituted for r_c and r_m in equations (16) and the resulting values of the R's are substituted into equation (8), the result is, to a good approximation,

$$v_2/v_g = \frac{\alpha R_L}{(r_e + r_b + R_g)[1 + R_L(1 + j\omega C_c r_c)/r_c] - ar_b} \tag{69}$$

The cutoff frequency f_{cc} is defined as the frequency at which the voltage across the load has dropped 3 db compared to its low-frequency value. This is the frequency at which the imaginary part of the denominator of (69) is equal to the real part. Solving for f_{cc} gives

$$f_{cc} = \frac{1}{2\pi C_c} \left[\frac{1}{R_L} + \frac{1}{r_c} - \frac{\alpha r_b}{R_L(r_e + r_b + R_g)} \right] \tag{70}$$

Substituting into this equation $C_c = 7(10)^{-12}$ farad, numerical values of the r's from (16), and the values $R_g = 91$ and $R_L = 4.58 \ (10)^6$ ohms, corresponding to maximum available gain gives

$$f_{cc} = 3390 \text{ cps.}$$

With these terminations, the low-frequency gain is 44.3 db. If R_g and R_L are reduced to 25 and 200,000 ohms, respectively, f_{cc} is raised to 23,500 cps and the gain is lowered to 37.2 db. A further reduction of R_L to 20,000 ohms increases f_{cc} to 0.22 megacycles and reduces the gain to 27.8 db. This corre-

sponds to a gain-bandwidth product of 1.2(10)^8 cps and shows that useful gain could be obtained at frequencies well above a megacycle, *provided* alpha and emitter cutoffs did not interfere.

Collector Cutoff in the Grounded-Emitter Stage

The procedure described in the last section leads, in this case, to

$$v_2/v_g = \frac{-R_L r_m/r_c + (R_L r_c/r_c)(1 + jr_c C_c \omega)}{r_e + (r_b + R_g)(1 - r_m/r_c) + [r_e R_L/r_c + (r_b + R_g)(r_e + R_L)/r_c](1 + jr_c C_c \omega)} \quad (71)$$

For the transistor of our numerical example, the imaginary term in the numerator is completely negligible at frequencies below (10)^9 cps. Neglecting it leads to

$$f_{cc} = \frac{1}{2\pi C_c} \frac{1 + R_L/r_c + [(r_b + R_g)/r_c][1 - (r_m - r_e - R_L)/r_c]}{R_L + [(r_b + R_g)/r_c](r_e + R_L)} \quad (72)$$

In this case the values of R_g and R_L (619 ohms and 671,000 ohms respectively) which correspond to maximum available gain give

$$f_{cc} = 3740 \text{ cps and}$$

$$\text{M.A.G.} = 53 \text{ db.}$$

Reducing R_L to 100,000 and increasing R_g to 1000 ohms gives

$$f_{cc} = 11{,}120 \text{ cps}$$

$$G = 50 \text{ db.}$$

For $R_g = 1000$ and $R_L = 10{,}000$,

$$f_{cc} = 97{,}900 \text{ cps}$$

$$G = 41.3 \text{ db.}$$

and for $R_g = R_L = 1000$ ohms,

$$f_{cc} = 943{,}000 \text{ cps}$$

$$G = 31.4 \text{ db.}$$

The gain-bandwidth product for this stage is 1.3(10)^9 cps as compared to 1.2(10)^8 cps for the same transistor connected as a grounded base amplifier. It should be pointed out, however, that this stage is particularly sensitive to change in α and on this account alpha cutoff may influence the response at fairly low frequencies. For example, when the terminating resistances are both 1000 ohms, reducing α from 0.9785 to 0.900 reduces the gain from 31.4 db to 0.2 db.

31

In this case

$$v_2/v_g = \cfrac{R_L}{[r_e + R_L + (r_b + R_g)(1 - r_m/r_c)] + (1/r_c)(r_b + R_g)(r_e + R_L)(1 + j\omega C_c r_c)} \qquad (73)$$

and

$$f_{cc} = \frac{1}{2\pi C_c}\left[\frac{1}{r_c} + \frac{1}{r_b + R_g} + \frac{1 - r_m/r_c}{r_e + R_L}\right] \qquad (74)$$

For matched impedances (R_g = 139,000 ohms and R_L = 2990 ohms),

$$f_{co} = 320{,}000 \text{ cps}$$

$$G = 16.7 \text{ db.}$$

The cutoff frequency can be raised by decreasing either R_g or R_L. With $R_g = 139{,}000$ and $R_L = 25$ ohms

$$f_{cc} = 9.77 \text{ megacycles}$$

$$G = 1.8 \text{ db}$$

The gain-bandwidth product in this case is $1.5(10)^7$.

NOISE

The data now available on noise are insufficient to give an adequate picture of the performance of n-p-n transistors in this respect. Such measurements as have been made, however, make it clear that these devices are very much quieter than early point-contact transistors reported on by Ryder and Kircher.

Transistor noise seems still to decrease with increasing frequency at a rate of something like 11 db per decade. It also decreases as the thickness of the p layer is decreased.

Of the order of half a dozen units of various dimensions have been measured at 1000 cps and have shown noise figures as low as 8 db and as high as 25 db.

The dependence of noise figure on operating point has been measured for only one transistor. As indicated in Fig. 28 and Fig. 29, these data show that the noise figure improves as V_c is reduced and that jt may be roughly independent of collector current. These data were taken on a grounded emitter stage with impedance match at the input terminals. Noise figure for this connection varies slightly with source impedance and has been found

to be a minimum when the source impedance is roughly equal to the input impedance of the stage.

It must be emphasized that this functional dependence of noise figure on operating point and source impedance has been measured for only one transistor. Further measurements may show that these results are not typical.
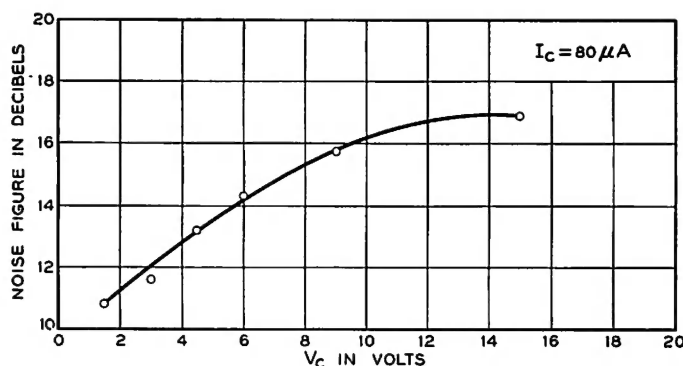

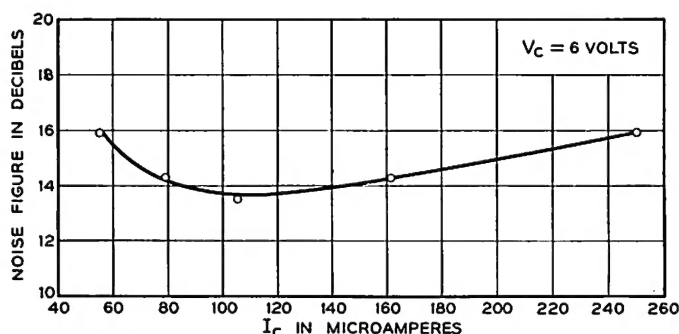Fig. 28—Noise figure increases with increasing collector voltage.

Fig. 29—Noise figure does not vary much with collector current.

FINAL COMMENTS

In this paper we have attempted to present what is known about the circuit performance of *n-p-n* transistors. Since these devices are still undergoing exploratory development and since only a limited number has been produced, it is obviously impossible to give statistical data on reproducibility or on such reliability factors as the effect of ambient temperature.

It is much too soon to know what properties may be achieved after further development, but the results obtained to date seem encouraging and worth reporting.

33

ACKNOWLEDGMENT

The authors are happy to acknowledge their indebtedness to W. Shockley who was first to conceive the n-p-n transistor and has provided much of the inspiration and guidance which has made its physical realization possible. His comments have been of great help in the preparation of this paper.

We are also much indebted to J. A. Morton for his encouragement and helpful guidance and to M. Sparks for providing most of the transistors which have been studied. We wish to thank L. O. Schott, L. C. Geiger, and K. D. Smith for taking some of the data presented and to thank G. Raisbeck and L. G. Schimpf for proofreading and correcting the manuscript.

REFERENCES

1. W. Shockley, "The Theory of p-n Junctions in Semiconductors and p-n Junction Transistors," *B. S. T. J.*, XXVIII, 435 (1949).
2. W. Shockley, "Electrons and Holes in Semiconductors," Van Nostrand (1950).
3. F. S. Goucher, G. L. Pearson, M. Sparks, G. K. Teal, W. Shockley, "Theory and Experiment for a Germanium p-n Junction," *Phys. Rev., 81*, 637 (1951).
4. W. Shockley, M. Sparks, G. K. Teal, "p-n Transistors," *Phys. Rev., 83*, 151 (1951).
5. R. M. Ryder, R. J. Kircher, "Some Circuit Aspects of the Transistor," *B. S. T. J.*, XXVIII, 367 (1949).

BELL TELEPHONE LABORATORIES, *incorporated*

463 West Street, New York 14, N. Y.